

CC Prostatic hyperplasia and to test environmental chemicals for
XX oestrogenic activity.

SO Sequence 485 AA:

Query Match 91.1%; Score 2554; DB 18; length 485;
Best local similarity 99.8%; from No. 2,76-235;
Matches 484; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 46 MITYSAVANNYSIPSVNTNIEGGPGKQTSPNVLWPTGHLSPLVHROLSHLYAEPOKS 105
DB 1 MITYSAVANNYSIPSVNTNIEGGPGKQTSPNVLWPTGHLSPLVHROLSHLYAEPOKS 60
QY 106 FWCEAKSLHETLVNRETKREKVSCHKCAKSPVAGSGKRAHCAVCSQYASGHTGWS 165
DB 61 FWCEAKSLHETLVNRETKREKVSCHKCAKSPVAGSGKRAHCAVCSQYASGHTGWS 120
QY 166 CHGCKAFKRSIOGHNDYICPAINQCTIKNFRKSCACPLPCYEVGWYCGSPPEPCG 225
DB 121 CHGCKAFKRSIOGHNDYICPAINQCTIKNFRKSCACPLPCYEVGWYCGSPPEPCG 180
QY 226 YHLYVRLAFAADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 285
DB 181 YHLYVRLAFAADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 240
QY 286 SAPFTASMMMSITKLAKDELVHMISWARKIGFVELSLPDQVRLLESGMVEVLMGLW 345
DB 241 SAPFTASMMMSITKLAKDELVHMISWARKIGFVELSLPDQVRLLESGMVEVLMGLW 300
QY 346 KSLDHGRKLEAFADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 405
DB 301 KSLDHGRKLEAFADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 360
QY 406 LNSSMPLVTATQDADSSKRLAHILNAVTDAVWYAKSGISSQOOSMTANLMLSHV 465
DB 361 LNSSMPLVTATQDADSSKRLAHILNAVTDAVWYAKSGISSQOOSMTANLMLSHV 420
QY 466 PHASNGKMEHLLNMKCKNVVIVYELLLHMLNANVLCNGNSSIIGSKAFAEDSKSKESQ 525
DB 421 PHASNGKMEHLLNMKCKNVVIVYELLLHMLNANVLCNGNSSIIGSKAFAEDSKSKESQ 480
QY 526 NFQSQ 530
DB 481 NFQSQ 485

```

RESULT 5

ID Y01597 standard; protein; 485 AA.

XX Y01597.

XX

DT 17 JUN 1999 (first entry)

XX

DE An estrogen receptor beta protein.

XX

KW differential estrogen receptor; ER-alpha; ER-beta; activation,
XX API site; therapy.

XX

OS Homo sapiens.

XX

PN W09911760-A1.

XX

PD 11-MAR-1999.

XX

PF 31-AUG-1998; 98W0-0518030.

XX

PK 04-SEP-1997; 97US-0923708.

XX

FA (REGG) UNIV CALIFORNIA.

XX

PI Gustafsson J., Kuiper GGM., Kushner PJ., Nilsson S;

P1 Paech K., Scanlan JS., Webb P.

XX
DR WP1: 1999-205173/17.
DR N-PSDB: X26686.

XX
PT Screening for test compounds which inhibit or activate estrogen
XX receptor beta (ER-beta) mediated activation at a API site - for use
XX in therapy

PS Claim 14: Page 43-44; 61pp; English.

XX
CC The present sequence represents an estrogen receptor-beta. The
CC specification describes a method for screening test compounds for
CC differential estrogen receptor (ER) alpha-mediated and ER beta-mediated
CC activation at an API site. The method is used for identifying compounds
XX which can be used in therapy.

SO Sequence 485 AA:

Query Match 90.4%; Score 2535; DB 20; length 485;
Best local similarity 99.2%; from No. 1,87-233;
Matches 481; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 46 MITYSAVANNYSIPSVNTNIEGGPGKQTSPNVLWPTGHLSPLVHROLSHLYAEPOKS 105
DB 1 MITYSAVANNYSIPSVNTNIEGGPGKQTSPNVLWPTGHLSPLVHROLSHLYAEPOKS 60
QY 106 FWCEAKSLHETLVNRETKREKVSCHKCAKSPVAGSGKRAHCAVCSQYASGHTGWS 165
DB 61 FWCEAKSLHETLVNRETKREKVSCHKCAKSPVAGSGKRAHCAVCSQYASGHTGWS 120
QY 166 CHGCKAFKRSIOGHNDYICPAINQCTIKNFRKSCACPLPCYEVGWYCGSPPEPCG 225
DB 121 CHGCKAFKRSIOGHNDYICPAINQCTIKNFRKSCACPLPCYEVGWYCGSPPEPCG 180
QY 226 YHLYVRLAFAADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 285
DB 181 YHLYVRLAFAADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 240
QY 286 SAPFTASMMMSITKLAKDELVHMISWARKIGFVELSLPDQVRLLESGMVEVLMGLW 345
DB 241 SAPFTASMMMSITKLAKDELVHMISWARKIGFVELSLPDQVRLLESGMVEVLMGLW 300
QY 346 KSLDHGRKLEAFADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 405
DB 301 KSLDHGRKLEAFADLVLDDECKEVEGTEIEHIMLALALSPBVLVLLLEAPPHVLSRP 360
QY 406 LNSSMPLVTATQDADSSKRLAHILNAVTDAVWYAKSGISSQOOSMTANLMLSHV 465
DB 361 LNSSMPLVTATQDADSSKRLAHILNAVTDAVWYAKSGISSQOOSMTANLMLSHV 420
QY 466 PHASNGKMEHLLNMKCKNVVIVYELLLHMLNANVLCNGNSSIIGSKAFAEDSKSKESQ 525
DB 421 PHASNGKMEHLLNMKCKNVVIVYELLLHMLNANVLCNGNSSIIGSKAFAEDSKSKESQ 480
QY 526 NFQSQ 530
DB 481 NFQSQ 485

```

RESULT 6

ID W33212 standard; protein; 477 AA.

XX W33212.

XX

AC W33212;

XX

DT 20-APR-1998 (first entry)

XX

DE human estrogen receptor protein.

XX

KW Oestrogen receptor protein; steroid; alternative splicing; estradiol;
XX estone; estril; screening.


```

XX Homo sapiens.
XX
XX E0798478-A2.
XX
XX 04-061-1997.
XX
XX 25-MAR-1997: 97EP-020090-A.
XX
XX 22-NV-1997: 96EP-02-0244.
XX
XX 28-MAR-1997: 96EP-020082-0.
XX
XX (AKU ) AKU ) MABEL NV.
XX
XX b1)Koma R. Mosselman S.
XX
XX WP1: 1997-474188/44.
XX
XX N (SDS) 186412.
XX
XX b1) DNA encoding estrogen receptor - useful in screening assay to
XX identify novel ligands of hormonal analogues.
XX
XX Claim 4: Page 18-20: 45pp: English.
XX
XX This sequence represents a novel estrogen binding protein. The cDNA
XX sequence which encodes this protein can be alternatively spliced
XX resulting in the detection of additional transcripts (see T86413).
XX This receptor is able to bind and be activated by estradiol, estrone and
XX estradiol, can be used in a screening assay for the identification of new
XX drugs and novel ligands of hormonal analogues.
XX
XX Sequence: 477 AA:

```

```

Query Match: 89.98% Score 2522: Db 18: Length 477:
Best Local Similarity: 100.00% Prod. No. 36732:
Matches: 477: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

```

```

01 54 MNVSPVNTNFRGCGQTSPNWKTPRHSPLVHQLSHLYAEFGKSPWCEAPSL 113
02 1 MYSIPSPVNTNFRGCGQTSPNWKTPRHSPLVHQLSHLYAEFGKSPWCEAPSL 113
03 144 PHILIPVNTNFRGCGQTSPNWKTPRHSPLVHQLSHLYAEFGKSPWCEAPSL 173
04 61 PHILIPVNTNFRGCGQTSPNWKTPRHSPLVHQLSHLYAEFGKSPWCEAPSL 126
05 124 KRSISGNNVYFPAINGTIFKNSKSGVAPEPVYVWYFGSPPEPGYPLVPKGP 233
06 121 KRSISGNNVYFPAINGTIFKNSKSGVAPEPVYVWYFGSPPEPGYPLVPKGP 180
07 244 SAGPGLHPCAGKAFSGTJAFPEVELLALSDPEGVITLLEAPPHVILSELSAETAS 293
08 161 SAGPGLHPCAGKAFSGTJAFPEVELLALSDPEGVITLLEAPPHVILSELSAETAS 240
09 294 MNMSTLNLALRFLVNTSMWAKRPGVPLSLFLWVLAESCMVEVLMMDIMWPSLDHPK 353
10 241 MNMSTLNLALRFLVNTSMWAKRPGVPLSLFLWVLAESCMVEVLMMDIMWPSLDHPK 300
11 444 ELPAGNVTLPKPKVWVSTIFETIMLTATSPPEPIKQHEFYLCVAMILLNSMYPL 413
12 401 ELPAGNVTLPKPKVWVSTIFETIMLTATSPPEPIKQHEFYLCVAMILLNSMYPL 360
13 414 VIATLALSSKRLAHLLNVATLAWVLAIAKSLSSGQSMKLANLMLSHVHNSNKM 473
14 401 VIATLALSSKRLAHLLNVATLAWVLAIAKSLSSGQSMKLANLMLSHVHNSNKM 420
15 474 EHLNLMKRNKVVVILLLEMLNAHLKCKSSITGSPKSPADSKSKSGSONPQSU 530
16 421 EHLNLMKRNKVVVILLLEMLNAHLKCKSSITGSPKSPADSKSKSGSONPQSU 477

```

```

10 W98125 standard; Protein: 549 AA:
XX
XX W98125:
XX
XX 21-JUN-1999 (first entry)
XX
XX Mouse estrogen receptor beta 1a (mER beta 1).
XX
XX estrogen receptor: ER beta 1; ER beta 2; mouse: 1020-21;
XX diagnostic: therapy: isoform.
XX
XX Mus musculus.
XX
XX W09605170-A1.
XX
XX 04-FEB-1999.
XX
XX 28-JUL-1998: 96MO-05155-09.
XX
XX 30-JUL-1997: 97OS-005421-0.
XX
XX 28-JUL-1997: 97OS-005485-0.
XX
XX (UYVA ) UNIV YALE.
XX
XX b1) Denton R. Fruston C. Gaffin R. Pascho-Paul G. M.
XX
XX WP1: 1999-142652/12.
XX
XX Human estrogen receptor beta 1 - useful in finding and modified for
XX tumors.
XX
XX Example 2: Fig 2a: 99pp: English.
XX
XX The present sequence is alternatively spliced isoform mER beta 1
XX of the murine complete estrogen receptor beta 1 (ER beta 1) of
XX mER beta 1). The mER beta 1 gene (see X2494) contains 5 exons,
XX including the newly described exon 7b, which is not found in ER
XX beta-1. Exon 5b is located in the third binding domain of mER
XX beta-1, and probably plays a standard role in mER beta-1 function.
XX The mER beta-1 clone was isolated by PCR of mRNA, and by
XX amplification from a mouse embryonic stem cell genome DNA library.
XX It lacks exon 5b. Preliminary data indicate that the mER beta 1
XX isoform may be more active than the full-length mER beta 1 (see
XX W98128). 2 other isoforms, mER beta 2 (see W98126) and rat ER
XX beta 4 (see W98127), have also been isolated. 1 and the human exon 5b
XX region (see X2494) has been identified, based on these sequences.
XX The invention provides methods for identifying mutants that block or
XX augment ER beta-1 mediated transcriptional regulation, methods for
XX determine whether ER beta-1 has isoforms, or to detect expressed in
XX tissues or cells, and methods of identifying and using mutants that
XX block the transcriptional regulation of genes by ER beta 1 or its
XX isoforms, which in turn modulates other physiological and pathological
XX processes. See, for fully expression of ER beta 1 is disclosed.
XX
XX Sequence: 549 AA:

```

```

Query Match: 89.28% Score 2522: Db 18: Length 549:
Best Local Similarity: 88.79% Prod. No. 36732:
Matches: 470: Conservative 27: Mismatches 25: Indels 0: Gaps 0:

```

```

01 1 MNMSTLNLALRFLVNTSMWAKRPGVPLSLFLWVLAESCMVEVLMMDIMWPSLDHPK 60
02 20 MNMSTLNLALRFLVNTSMWAKRPGVPLSLFLWVLAESCMVEVLMMDIMWPSLDHPK 79
03 61 NVINLEKSGMGTLSINVLVWVNTSPVYVWYFGSPPEPGYPLVPKGP 120
04 60 NVINLEKSGMGTLSINVLVWVNTSPVYVWYFGSPPEPGYPLVPKGP 119
05 121 PHLFPAVSNPASFVITLSSKRLAHLLNVATLAWVLAIAKSLSSGQSMKLANLMLSHVHNSNKM 190
06 140 PHLFPAVSNPASFVITLSSKRLAHLLNVATLAWVLAIAKSLSSGQSMKLANLMLSHVHNSNKM 109

```


D6	329	Iodotyrosine:methylglucosyl glucuronide	379
C7	VLP	PRFQKCEVCLILFLITMIALTSPPFEKLGRKYGVKA	403
D6	380	Virus-coded protein kinase encoded by cytomegalovirus strain 1B-1; also referred to as gamma-cyprinid herpesvirus 1 UL130 protein	439
OY	403	MILINSSWPIVTATODDSSRKLAHLINAVTDALWWIAKSIGISSOOSMLANLIML	462
D6	440	mollusmyofibrin-associated protein 1-like domain-containing protein 1 (MyoP1); also known as myofibrin-associated protein 1 (MyoP1)	499
OY	462	SVHVASPCDHPTTPGEEFWVVYLLEELNANHIFDKSSTISDESFAEESKE	522
D6	500	Schistosome blood-liver-stage-specific antigen 1 (SLA-1); also known as Schistosoma mansoni SLA-1	553
OY	523 G 523 I		
D6	560 g 560		
	SEQUENT 1: WI4723		
ID	WI4723 standard J, Protein, 425 AA.		
XX	WI4723:		
AC	WI4723:		
D1	08 JUN 1997 (first entry)		
XX			
DE	Fat oestrogen receptor beta (fcr beta).		
XX			
KW	Orphan receptor, oestrogen receptor beta, ER-beta, nuclear receptor, prostate cancer benign prostatic hyperplasia; early onset, endocrine disease, osteoporosis, environment pollutant.		
BA	Rattus sp.		
XX			
OS			
FI	Key location/qualifiers		
EH	Domain 104..169		
EI	/label=DNA-binding_domain		
FI	Domain 259..457		
FI	/label=Ligand-binding_domain		
PX	W09709448-AZ.		
PN			
XX			
PD	14-MAR-1997.		
XX			
PF	09-SEP-1996;	96MO-Ep03933.	
XX			
PR	08-MAY-1996;	96GB-0009576.	
PR	08-SEP-1995;	95GB-0018272.	
PR	15-MAR-1996;	96GB-0005550.	
PK	11-APR-1996;	96GB-0007582.	
XX	(KARO-) KARO BIO AB.		
PA			
XX			
PI	Emmarik E, Gustafsson O, Kulper GG;		
XX	MP1:1997-192842/17.		
D6	N-FSHB: T92842.		
XX			
PI	New isolated oestrogen receptor beta - used to develop prods. for treating e.g. cancers, CNS disorders, osteoporosis or cardiovascular disease		
XX			
PS	claim 1: Fig 1: 45pp: English.		
XX			
OC	A novel rat orphan receptor (WI4723) is related to the known oestrogen receptor ER-alpha, and has been designated ER-beta. It is an oestrogen receptor related nuclear receptor. The ER-beta amino acid sequence was deduced from a cDNA clone (T92842) isolated from a rat prostate cDNA library. Rat, human and mouse Er-beta (WI4723-25) can be used to isolate molecules for use in the treatment of cardiovascular diseases, central nervous system		

[illegible]

01 Unpublished. Kulper GDM, Kusinor PJ, Nilsson S.
 02 Jacob K, Scudiero JS, Webb P.
 03
 04 WPI: 1997-205174/17.
 05 N-PSDB: X2665.
 06
 07 Screening for test compounds which inhibit or activate estrogen
 08 receptor beta (ER-beta) mediated activation at a AP1 site - for use
 09 in therapy.
 10
 11 Main: 41-43; 44pp: English.
 12
 13 The present sequence represents an estrogen receptor beta. The
 14 specification describes a method for screening test compounds for
 15 differential estrogen receptor (ER) alpha-mediated and ER-beta-mediated
 16 activation at an AP1 site. The method is used for identifying compounds
 17 which can be used in therapy.
 18
 19 Sequence 485 AA:
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

01 nuclear receptor: prostate cancer: beta1n: prostate: ER-beta: nuclear
 02 ovary: cancer: endovascular: disease: estrogen: ER-beta: nuclear
 03 environmentally: pediatric:
 04
 05 Mus sp.
 06
 07
 08
 09
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520

QY 445 PSIDHPEELAEGLVETCEKCEVEELLEELMLALISPEERKIQHREYVAKAM 455
 DB 401 FADHGFVETGPELVCHDQKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 390
 QY 406 INSSNVIATVATPAASERKARIINAVTAVVWVARSQSSQSSQSSQSSQSSQSS 455
 DB 361 INSSNVIATVATPAASERKARIINAVTAVVWVARSQSSQSSQSSQSSQSSQSS 420
 QY 456 FHAEHPEELAEGLVETCEKCEVEELLEELMLALISPEERKIQHREYVAKAM 505
 DB 421 FHAEHPEELAEGLVETCEKCEVEELLEELMLALISPEERKIQHREYVAKAM 480
 QY 526 NPGSQ 530
 DB 481 NPGSQ 485

RESULT 14
 Y04435
 ID Y04435 standard: Protein: 510 AA.

XX Y04435:

DE 25-JUN-1999 (first entry)

XX Murine mER beta-2 clone protein sequence.

XX Estrogen receptor beta: estrogen receptor beta; ER beta.

XX Mus SP.

XX W09905171-A1.

DB 04-FEB-1999.

PE 28-JUL-1998: 98W0-US155410.

PE 30-JUL-1997: 97US-0054210.

PE 28-JUL-1997: 97US-0053869.

PA (UUYA) UNIV YALF.

PI Baron R, Chambon P, Denton K, Dupont S, Garnier J;

PT Novel murine estrogen receptor-beta genes and splice variants -

PT useful for treating estrogen receptor beta mediated disorders

PS claim 9; Fig 2b; 89pp; English.

CC The present invention describes the murine complete estrogen receptor

CC beta (mER beta) also designated mER beta-3). The present sequence

CC can be used to screen for agents that modulate or block the interaction

CC between the ER beta c protein and physiological, stilbene or

CC triphenylethylene (anti)estrogens. The antibody which binds the protein

CC can be used in assays to detect activation of the ER beta protein and

CC can be used for subcellular localization of the protein. The

CC complementary nucleic acid can be used to inhibit expression of the

CC ER beta c nucleic acid.

XX Sequence 510 AA;

Query Match: #0 14; Score 2245.5; DB 20; Length 510;
 Best Local Similarity: 91.4%; Prod. No: 9126-206;
 Matches 417; Conservative 25; Mismatches 35; Indels 39; Gaps 1.

QY 1 MERNKPSINSGSYNCQSLPLMHOSITYFSSVDSHMYAMFYFAVMNYSIFS 60
 DB 20 MERNKPSINSGSYNCQSLPLMHOSITYFSSVDSHMYAMFYFAVMNYSIFS 79

QY 41 NVNLEPQV 120
 DB 20 NVNLEPQV 139
 QY 173 FEFVVSQV 180
 DB 140 FEFVVSQV 197
 QY 181 NVNLEPQV 240
 DB 158 NVNLEPQV 220
 QY 241 CACAKESQCHAPVPELTPALSPEDVITLPAEPHVVISPSAPFASMMMSLTK 300
 DB 221 CACAKESQCHAPVPELTPALSPEDVITLPAEPHVVISPSAPFASMMMSLTK 290
 QY 301 IADKPIVHMSWARKPEPEVLSIFQVVAIPSSQWVIMKIMKESIDHDKSLFAPDL 360
 DB 281 IADKPIVHMSWARKPEPEVLSIFQVVAIPSSQWVIMKIMKESIDHDKSLFAPDL 340
 QY 361 VLDPEQGVQV 420
 DB 341 VLDPEQGVQV 400
 QY 421 DESPELAEINAVTAVVWVARSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSS 480
 DB 401 DESPELAEINAVTAVVWVARSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSS 460
 QY 481 GYNNVWVYVATPAASERKARIINAVTAVVWVARSQSSQSSQSSQSSQSSQSSQSSQSS 540
 DB 461 GYNNVWVYVATPAASERKARIINAVTAVVWVARSQSSQSSQSSQSSQSSQSSQSSQSS 510

RESULT 15
 W98126
 ID W98126 standard: Protein: 510 AA.

XX W98126:

DE 21-JUN-1999 (first entry)

XX Mouse estrogen receptor beta isoform mER beta-2.

XX Estrogen receptor; ER beta-2; ER beta-c; mouse; tumor;

XX diagnosis; therapy; isoform.

XX Mus musculus.

PN W09905170-A1.

PD 04-FEB-1999.

PE 28-JUL-1998: 98W0-US155410.

PE 30-JUL-1997: 97US-0054210.

PE 28-JUL-1997: 97US-0053869.

PA (UUYA) UNIV YALF.

PI Denton K, Fraslon C, Galien K, Kische-Kron M;

PT Human estrogen receptor beta-3 useful in treating and monitoring

PT tumors

PS Example 2; Fig 2b; 89pp; English.

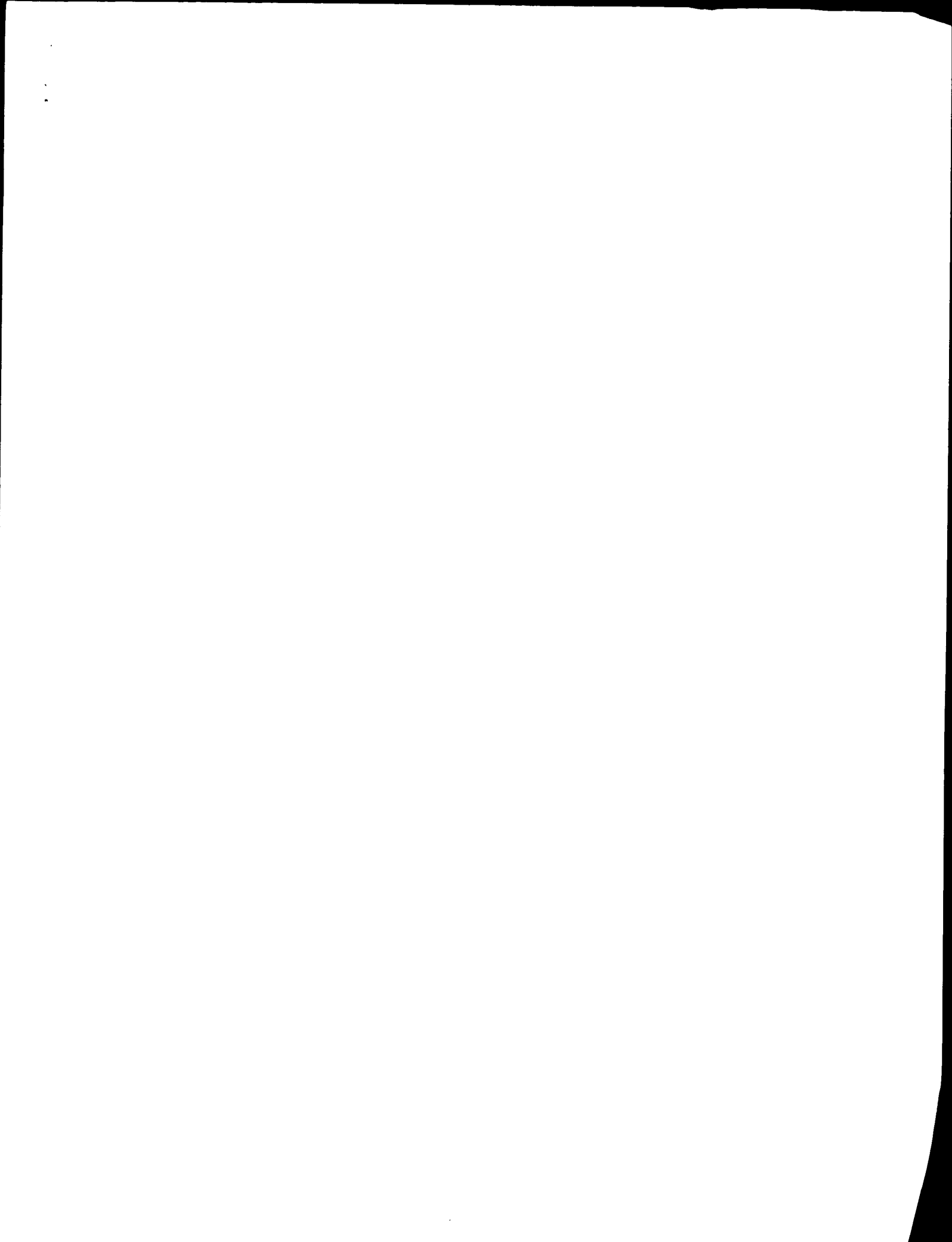
CC The present sequence is alternatively spliced isoform mER beta 2

CC of the murine complete estrogen receptor-beta (ER beta-c, or

CC mER beta-3). The mER beta 3 gene (see X24435) contains 9 exons,

CC including the coding exon 5b, which is not found in ER

CC beta-1. Exon 5b is located in the ligand binding domain of mER



Sanctore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

on protein protein search using SW model

run on: April 19, 2001, 06:29:21 : Search time 51.46 seconds

(without alignments)
452,806 Million cell updates/sec

11700
protein score: 2805
Sequence: 1 MATRNSPSSNPSNPNQSD ESNAPSKSKESQNSQNSQ 530

Search factors:
BLASTNM62
Gapop: 10.0 : Gapext: 0.5

Searched: 3445 seqs, 4425540 residues

Total number of hits satisfied chosen parameters: 93435

Minimum hit seq length: 0

Maximum hit seq length: 200000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database: 1 SwissProt_493*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2805	100.0	530	1	ESR2_HUMAN	Q92731 homo sapien
2	2615	98.2	530	1	ESR2_FALSA	Q96171 callitrix
3	2496	89.0	530	1	ESR2_RAI	Q62986 rattus norv
4	2496	88.9	530	1	ESR2_MOUSE	Q08537 mus muscul
5	2453.5	87.8	527	1	ESR2_B_VIR	Q94845 bos taurus
6	2453.5	87.5	527	1	ESR2_SHEEP	Q91115 ovis aries
7	2401	74.9	524	1	ESR2_STUVO	Q99022 sturnus vul
8	1969	69.8	472	1	ESR2_D21JA	Q94511 coturnix co
9	1948.5	69.5	472	1	ESR2_CHICK	Q99025 gallus gall
10	1649	55.2	573	1	ESR2_ANGOA	Q14012 anquilla ja
11	1466	52.3	538	1	ESR2_TAKAI	Q96659 carassius a
12	1440.5	51.4	538	1	ESR2_NCMY	P57782 oncorhynch
13	1441	51.0	673	1	ESR2_MTCIN	P57781 microtyph
14	1404	49.8	279	1	ESR2_MACMO	Q91165 macaca mula
15	1385.5	49.4	559	1	ESR2_SPAUD	Q94602 sparus auri
16	1381	49.2	575	1	ESR2_ICIPI	Q91431 ictalurus p
17	1378.5	49.1	611	1	ESR2_CAKAU	Q91419 carassius a
18	1373	48.9	565	1	ESR2_MTCIN	P57783 microtyph
19	1362	48.2	557	1	ESR2_BERN	Q93432 creochromis
20	1283.5	44.0	595	1	ESR1_HUMAN	P03472 homo sapien
21	1283.5	43.8	587	1	ESR1_P96B0	Q91250 poephilia qu
22	1275.5	43.7	589	1	ESR1_CHICK	P06212 gallus gall
23	1270	43.5	594	1	ESR1_BORSE	Q91098 equus cabal
24	1219.5	43.5	595	1	ESR1_PIG	Q29640 sus scrofa
25	1219	43.5	595	1	ESR1_MESAU	Q94315 mesocricetu
26	1218.5	43.4	599	1	ESR1_MOUSE	P19785 mus muscul
27	1216.5	43.4	595	1	ESR1_RAI	P06211 rattus norv
28	1214	43.3	585	1	ESR1_BERN	Q93433 creochromis
29	1214	43.3	585	1	ESR1_XENLA	P81559 xenopus lae
30	1204	42.9	569	1	ESR1_BRAVE	P57717 brachydanio
31	1204	42.9	579	1	ESR1_SPAUD	Q94602 sparus auri
32	1199	42.7	560	1	ESR1_ORFA	P50241 oryzias lat
33	1194.5	42.6	561	1	ESR1_TAGMA	Q42132 paiprus majo

34	1184.5	42.2	563	1	ESR1_BRAVE	P57717 brachydanio
35	1177	42.0	525	1	ESR1_MTCIN	P57783 microtyph
36	1165	41.5	622	1	ESR1_NCMY	P57782 oncorhynch
37	1160.5	41.4	617	1	ESR1_ICIPI	Q91431 ictalurus p
38	1144.5	40.4	555	1	ESR1_SAIJA	P06212 gallus gall
39	877	41.3	507	1	ESR1_BERN	Q93432 creochromis
40	799.5	28.5	549	1	ESR1_ANGOA	Q14012 anquilla ja
41	744.5	26.5	193	1	ESR2_PIG	Q94315 mesocricetu
42	648	23.1	498	1	ESR3_HUMAN	Q944 homo sapien
43	648	23.1	498	1	ESR3_MOUSE	Q9113 mus muscul
44	641.5	22.5	433	1	ESR2_RAI	P06211 rattus norv
45	640	22.5	500	1	ESR2_HUMAN	Q92731 homo sapien

ALIGNMENTS

Result	ID	Sequence	Score	Length	DB	ID	Description
1	ESR2_HUMAN	Q92731	100.0	530	1	ESR2_HUMAN	Q92731 homo sapien
2	ESR2_HUMAN	Q92731	98.2	530	1	ESR2_FALSA	Q96171 callitrix
3	ESR2_HUMAN	Q92731	89.0	530	1	ESR2_RAI	Q62986 rattus norv
4	ESR2_HUMAN	Q92731	88.9	530	1	ESR2_MOUSE	Q08537 mus muscul
5	ESR2_HUMAN	Q92731	87.8	527	1	ESR2_B_VIR	Q94845 bos taurus
6	ESR2_HUMAN	Q92731	87.5	527	1	ESR2_SHEEP	Q91115 ovis aries
7	ESR2_HUMAN	Q92731	74.9	524	1	ESR2_STUVO	Q99022 sturnus vul
8	ESR2_HUMAN	Q92731	69.8	472	1	ESR2_D21JA	Q94511 coturnix co
9	ESR2_HUMAN	Q92731	69.5	472	1	ESR2_CHICK	Q99025 gallus gall
10	ESR2_HUMAN	Q92731	55.2	573	1	ESR2_ANGOA	Q14012 anquilla ja
11	ESR2_HUMAN	Q92731	52.3	538	1	ESR2_TAKAI	Q96659 carassius a
12	ESR2_HUMAN	Q92731	51.4	538	1	ESR2_NCMY	P57782 oncorhynch
13	ESR2_HUMAN	Q92731	51.0	673	1	ESR2_MTCIN	P57781 microtyph
14	ESR2_HUMAN	Q92731	49.8	279	1	ESR2_MACMO	Q91165 macaca mula
15	ESR2_HUMAN	Q92731	49.4	559	1	ESR2_SPAUD	Q94602 sparus auri
16	ESR2_HUMAN	Q92731	49.2	575	1	ESR2_ICIPI	Q91431 ictalurus p
17	ESR2_HUMAN	Q92731	49.1	611	1	ESR2_CAKAU	Q91419 carassius a
18	ESR2_HUMAN	Q92731	48.9	565	1	ESR2_MTCIN	P57783 microtyph
19	ESR2_HUMAN	Q92731	48.2	557	1	ESR2_BERN	Q93432 creochromis
20	ESR2_HUMAN	Q92731	44.0	595	1	ESR1_HUMAN	P03472 homo sapien
21	ESR2_HUMAN	Q92731	43.8	587	1	ESR1_P96B0	Q91250 poephilia qu
22	ESR2_HUMAN	Q92731	43.7	589	1	ESR1_CHICK	P06212 gallus gall
23	ESR2_HUMAN	Q92731	43.5	594	1	ESR1_BORSE	Q91098 equus cabal
24	ESR2_HUMAN	Q92731	43.5	595	1	ESR1_PIG	Q29640 sus scrofa
25	ESR2_HUMAN	Q92731	43.5	595	1	ESR1_MESAU	Q94315 mesocricetu
26	ESR2_HUMAN	Q92731	43.4	599	1	ESR1_MOUSE	P19785 mus muscul
27	ESR2_HUMAN	Q92731	43.4	595	1	ESR1_RAI	P06211 rattus norv
28	ESR2_HUMAN	Q92731	43.3	585	1	ESR1_BERN	Q93433 creochromis
29	ESR2_HUMAN	Q92731	43.3	585	1	ESR1_XENLA	P81559 xenopus lae
30	ESR2_HUMAN	Q92731	42.9	569	1	ESR1_BRAVE	P57717 brachydanio
31	ESR2_HUMAN	Q92731	42.9	579	1	ESR1_SPAUD	Q94602 sparus auri
32	ESR2_HUMAN	Q92731	42.7	560	1	ESR1_ORFA	P50241 oryzias lat
33	ESR2_HUMAN	Q92731	42.6	561	1	ESR1_TAGMA	Q42132 paiprus majo

[illegible]

```

|||||
ID 1 LADELVHMSWAKRIQGVVLSLFFVRLFSWMVYMMQIMKSLDHQKLLFAPDL 360
401 VILQEPKRVHETTFETFMATTSFPELRLEQREYLVKAMILLNSMYPVATVGA 420
|||||
406 VILQEPKRVHETTFETFMATTSFPELRLEQREYLVKAMILLNSMYPVATVGA 420
421 DSSKRIAMINAVTDAVWVAVASSTSSVQVSMPLANMLLSIVPASKKMHLLNMK 480
|||||
426 DSSKRIAMINAVTDAVWVAVASSTSSVQVSMPLANMLLSIVPASKKMHLLNMK 480
441 GKNVAVVHLLLEMLNAHVLKPTKSSITGSGSPASKSKKPSKPSQVSO 540
441 GKNVAVVHLLLEMLNAHVLKPTKSSITGSGSPASKSKKPSKPSQVSO 540

RESULT 2
ESR2_RAI1A STANDARD PR1 530 AA
Q95171
01-OCT-2000 (rel. 40) (created)
01-OCT-2000 (rel. 40) (last sequence update)
01-OCT-2000 (rel. 40) (last annotation update)
1E ESTROGEN RECEPTOR BETA (ER-BETA);
2N ESR2 OR NR6A2;
3S CALITRUX LACTONIS (Common marmoset);
4P Eukaryotic Motif: Steroid Hormone Receptor;
5M Mammalian Endocrine Primates: Platyrrhini: Callitrichidae: Callitrix
6X Nucleotide: 484
7N SEQUENCE FROM N.A.
8P TISSUE TESTES;
9L SUBMITTER (N.A.): 1999 (to the EMBL/GenBank/DDBJ databases);
10 1 FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
11 ER-ALPHA AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
12 ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER;
13 1 SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
14 ALPHA (BY SIMILARITY)
15 1 SUBCELLULAR LOCATION: NUCLEAR.
16 1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
17 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
18 1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
19 NR4 SUBFAMILY.

This SwissProt entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
of the European Bioinformatics Institute. There are no restrictions on its
use by non profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcements/
or send an email to license@isb-sib.ch).

EMBL: Y09727, ZAT0514.2;
DB HSSP: P04722; HRP;
DB InterPro: IPR000524;
DB InterPro: IPR000546;
DB InterPro: IPR016286;
DB InterPro: IPR001724;
DB Pfam: PF00104; hormone_1;
DB Pfam: PF00105; ZF-HR;
DB PRINTS: PR00047; STEROIDRENER;
DB PRINTS: PR00450; VITAMINR;
DB PRINTS: PR00498; STER0486NER;
DB ERSITE: PR00041; NUCLEAR RECEPTOR;
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Phosphorylation;
KW ZINC_FINGER 1 148 MODULATING;
F1 LNA_BIND 149 214 NUCLEAR RECEPTOR-TYPE;
F1 ZN_FING 149 165 C4-TYPE;
F1 ZN_FING 185 209 C4-TYPE;
F1 DOMAIN 215 530 STEROID-BINDING;
F1 M0_KICS 87 87 PHOSPHORYLATION (BY SIMILARITY).

```

```

F1 M0_KICS 484 488 PHOSPHORYLATION (BY SIMILARITY);
SD SEQUENCE 530 AA: 53087 MW: 43499.1/DAA: 3.461/0.604.

Query Match 93.4% Score 401.1 E-160 Length 530
Pos: Local Similarity 92.4% Pos: Local Similarity 92.4%
Matches 4897 Conserved 201 Mismatches 240 Gaps 0

QY 1 MDRKSNLSINSSVNPQSLTLEHREYLPSSVESHHHYAMIPYAMVNSHS 60
|||||
DB 1 MDRKSNLSINSSVNPQSLTLEHREYLPSSVESHHHYAMIPYAMVNSHS 60
QY 61 NVNLEEGPQSLTINVMKPTKHLSPVVRGLSHYATVKSIVTAPSTHLLAV 120
|||||
DB 61 NVNLEEGPQSLTINVMKPTKHLSPVVRGLSHYATVKSIVTAPSTHLLAV 120
QY 121 BELLEPEVSNEASVLEKSKALAEAYVYAKAGAWVDEKAKENSGH 180
|||||
DB 121 BELLEPEVSNEASVLEKSKALAEAYVYAKAGAWVDEKAKENSGH 180
QY 123 PTFEPVVSNRASVTVKSPKQAMVAVSQVAVGVWVWSVQVKAPEKSSQH 190
|||||
DB 123 PTFEPVVSNRASVTVKSPKQAMVAVSQVAVGVWVWSVQVKAPEKSSQH 190
QY 191 RNYVETALNLTTPRPPSCVAVPLRPVYVYKVVQVPEPQVYEVVPSALQ 240
|||||
DB 191 RNYVETALNLTTPRPPSCVAVPLRPVYVYKVVQVPEPQVYEVVPSALQ 240
QY 241 GAKKAPSGSHAPVPHLLHATSEPVVLLHAEHVLSHSAHIFASMMSTIK 300
|||||
DB 241 GAKKAPSGSHAPVPHLLHATSEPVVLLHAEHVLSHSAHIFASMMSTIK 300
QY 401 LADELVHMSWAKRIQGVVLSLFFVRLFSWMVYMMQIMKSLDHQKLLFAPDL 460
|||||
DB 401 LADELVHMSWAKRIQGVVLSLFFVRLFSWMVYMMQIMKSLDHQKLLFAPDL 460
QY 461 VILQEPKRVHETTFETFMATTSFPELRLEQREYLVKAMILLNSMYPVATVGA 480
|||||
DB 461 VILQEPKRVHETTFETFMATTSFPELRLEQREYLVKAMILLNSMYPVATVGA 480
QY 481 GKNVAVVHLLLEMLNAHVLKPTKSSITGSGSPASKSKKPSKPSQVSO 540
|||||
DB 481 GKNVAVVHLLLEMLNAHVLKPTKSSITGSGSPASKSKKPSKPSQVSO 540

RESULT 3
ESR2_RAI1 STANDARD PR1 530 AA:
Q95171
01-OCT-2000 (rel. 40) (created)
01-OCT-2000 (rel. 40) (last sequence update)
01-OCT-2000 (rel. 40) (last annotation update)
1E ESTROGEN RECEPTOR BETA (ER-BETA);
2N ESR2 OR NR6A2;
3S CALITRUX LACTONIS (Common marmoset);
4P Eukaryotic Motif: Steroid Hormone Receptor;
5M Mammalian Endocrine Primates: Platyrrhini: Callitrichidae: Callitrix
6X Nucleotide: 484
7N SEQUENCE FROM N.A. (ISOPRM BETA-2);
8P STRAIN SPRAGUE-DAWLEY; TISSUE: Prostate;
9L SUBMITTER (N.A.): 1999 (to the EMBL/GenBank/DDBJ databases);
10 1 FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
11 ER-ALPHA AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
12 ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER;
13 1 SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
14 ALPHA (BY SIMILARITY)
15 1 SUBCELLULAR LOCATION: NUCLEAR.
16 1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
17 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
18 1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
19 NR4 SUBFAMILY.

This SwissProt entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
of the European Bioinformatics Institute. There are no restrictions on its
use by non profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcements/
or send an email to license@isb-sib.ch).

EMBL: Y09727, ZAT0514.2;
DB HSSP: P04722; HRP;
DB InterPro: IPR000524;
DB InterPro: IPR000546;
DB InterPro: IPR016286;
DB InterPro: IPR001724;
DB Pfam: PF00104; hormone_1;
DB Pfam: PF00105; ZF-HR;
DB PRINTS: PR00047; STEROIDRENER;
DB PRINTS: PR00450; VITAMINR;
DB PRINTS: PR00498; STER0486NER;
DB ERSITE: PR00041; NUCLEAR RECEPTOR;
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Phosphorylation;
KW ZINC_FINGER 1 148 MODULATING;
F1 LNA_BIND 149 214 NUCLEAR RECEPTOR-TYPE;
F1 ZN_FING 149 165 C4-TYPE;
F1 ZN_FING 185 209 C4-TYPE;
F1 DOMAIN 215 530 STEROID-BINDING;
F1 M0_KICS 87 87 PHOSPHORYLATION (BY SIMILARITY).

```



```

OY 401 LAKEELVHMI SMARK I PGVEEL SI PGVVL ELSOMVLIMLIMKPSIDHPKLI PAAPL 360
DB 401 LAKEELVHMI SMARK I PGVEEL SI PGVVL ELSOMVLIMLIMKPSIDHPKLI PAAPL 360
OY 461 VLDDEKCEVGLI I FEMILATTAVERELI LQHELELVKAMILLNSMYLATASGEA 420
DB 461 VLDDEKCEVGLI I FEMILATTAVERELI LQHELELVKAMILLNSMYLATASGEA 420
OY 421 LSEKELMLLNAVITAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI 189
DB 421 LSEKELMLLNAVITAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI 189
OY 481 CRNVVAVYDILMLLNAHVLIPKSSITGSECHAEKSKSPKSPQSO 530
DB 481 CRNVVAVYDILMLLNAHVLIPKSSITGSECHAEKSKSPKSPQSO 530
DB 481 CRNVVAVYDILMLLNAHVLIPKSSITGSECHAEKSKSPKSPQSO 530

RESULT 5
ESR2_SHEEP STANDARD: PRT: 527 AA.
ID ESR2_SHEEP
AC 09J015: 09J015:
DT 01-OCT-2000 (rel. 40, last sequence update)
DE ESTROGEN RECEPTOR BETA (ER-BETA).
GN ESR2 OR NR1A2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Artiodactyla; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bovinae; Bos.
OC NCBI_TaxID: 9913.
RN 111
RP SEQUENCE FROM R.A.
RC STRAIN-HOLSTEIN; TISSUE-Ovarian follicle;
RX MEDLINE: 99150196; PubMed: 10026117;
RA Rosenfeld C.S., Yuan X., Manikam M., Calder M.D., Garverick H.A.,
RA Ludman D.B.;
R1 "Cloning, sequencing, and localization of bovine estrogen receptor-
R1 beta within the ovarian follicle."
R1 J. Biol. Reprod. 60:691-697(1999).
RN 121
RP SEQUENCE FROM N.A.
RC Strain: 9934179; Tissue: Follicle;
RA Mather N., Houtas C., Tillmann G., Twell R.;
R1 "Cloning of bovine estrogen receptor beta (Erbeta): expression of
R1 novel deleted isoforms in reproductive tissues."
R1 Mol. Cell. Endocrinol. 152:37-45(1999).
CC 1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
CC ER ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
CC MAY PLAY A ROLE IN OVARIAN FOLLICULAR GROWTH AND MATURATION.
CC 1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
CC ALPHA (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: NUCLEAR.
CC 1- TISSUE SPECIFICITY: PRESENT IN GRANULOSA CELLS OF ANTRAL FOLLICLES
CC IN VARIOUS STAGES OF FOLLICULAR GROWTH.
CC 1- DOMAIN: COMPOSED OF LIGAND BINDING, A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC 1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NR1 SUBFAMILY.
CC -----
CC This Swiss Prot entry is copyright. It is produced through a collaboration
CC between the European Bioinformatics Institute and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/infocentre/
CC or send an email to license@ebi.ac.uk).
CC -----
DB EMBL: AF110402; AAC24432.1;
DB EMBL: Y18017; CA653661.1; ALL_INIT.
DB HSCB: F03372; HSCB.

```

```

DB InterPro: IPR000324; -
DB InterPro: IPR000536; -
DB InterPro: IPR001628; -
DB InterPro: IPR001723; -
DB InterPro: IPR001728; -
DB Pfam: PF00104; hormone_rec_1.
DB Pfam: PF00105; 2f-04; 1.
DB PRINTS: PR00047; STR61P1RINGER.
DB PRINTS: PR00350; VITAMINDR.
DB PRINTS: PR00398; STERIODHOMR.
DB PRINTS: PR00546; THYROIDHOMR.
DB PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
DB PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
DB Zinc finger, steroid binding; Phosphorylation.
KW Receptor; Transcription regulation; DNA binding; Nuclear protein;
KW Zinc finger, steroid binding; Phosphorylation.
FT DOMAIN 1 145
FT DNA_BIND 146 211
FT ZN_FING 146 166
FT ZN_FING 182 206
FT DOMAIN 212 527
FT MOD_RES 84 84
FT MOD_RES 485 485
FT CONFLICT 60 60
FT SEQUENCE 527 AA; 59031 MW; 9CFEELD6F4ACB4 CRC64;
SO QUERY MATCH
Query Match 87.8%; Score 2463.5; DH 1; Length 527;
Best Local Similarity 87.9%; Pred. No. 3,56-185;
Matches 466; Conservative 22; Mismatches 349; Indels 4; Gaps 1;
OY 1 MKIKNSPSSINSSSYNGQSILPLEHGSIVLHSSVSHIHIVYAMFYSPAVNYSIPS 60
DB 1 MKYKNSPSSINSSSYNGQSILPLEHGSIVLHSSVSHIHIVYAMFYSPAVNYSIPS 60
OY 61 NVINIGSPSPVITSPNVIPVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIV 120
DB 61 NVINIGSPSPVITSPNVIPVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIV 117
OY 121 RETIKKSVNGSPSPVITSPNVIPVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPI 180
DB 121 RETIKKSVNGSPSPVITSPNVIPVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPI 177
OY 181 NDYLOVANGSPSPVITSPNVIPVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPI 240
DB 181 NDYLOVANGSPSPVITSPNVIPVPIVPIVPIVPIVPIVPIVPIVPIVPIVPIVPI 237
OY 241 GAKKAPSGOHAPVPELILDALSPROVLITLLEAPPVILSRSPAPTEASMMSTLK 300
DB 241 GAKKAPSGOHAPVPELILDALSPROVLITLLEAPPVILSRSPAPTEASMMSTLK 297
OY 238 GSKTRNGPMTKRVKVELLSALSPQVLITLLEAPPVILSRSPAPTEASMMSTLK 297
DB 238 GSKTRNGPMTKRVKVELLSALSPQVLITLLEAPPVILSRSPAPTEASMMSTLK 297
OY 301 LAKEELVHMI SMARK I PGVEEL SI PGVVL ELSOMVLIMLIMKPSIDHPKLI PAAPL 460
DB 301 LAKEELVHMI SMARK I PGVEEL SI PGVVL ELSOMVLIMLIMKPSIDHPKLI PAAPL 457
OY 361 VLDDEKCEVGLI I FEMILATTAVERELI LQHELELVKAMILLNSMYLATASGEA 420
DB 361 VLDDEKCEVGLI I FEMILATTAVERELI LQHELELVKAMILLNSMYLATASGEA 417
OY 358 VLDDEKCEVGLI I FEMILATTAVERELI LQHELELVKAMILLNSMYLATASGEA 417
DB 358 VLDDEKCEVGLI I FEMILATTAVERELI LQHELELVKAMILLNSMYLATASGEA 417
OY 421 LSEKELMLLNAVITAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI 480
DB 421 LSEKELMLLNAVITAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI VAVI 477
OY 481 CRNVVAVYDILMLLNAHVLIPKSSITGSECHAEKSKSPKSPQSO 530
DB 481 CRNVVAVYDILMLLNAHVLIPKSSITGSECHAEKSKSPKSPQSO 527
DB 478 CRNVVAVYDILMLLNAHVLIPKSSITGSECHAEKSKSPKSPQSO 527

```


DR	EMIL1: A0004356; HAAL19851.1; -	
DR	HSSP: F03372; HSCP	
DR	InterPro: IPR000536; -	
DR	InterPro: IPR001628; -	
DR	InterPro: IPR001723; -	
DR	PIfam: PF00104; hormone_rec; 1.	
DR	PIfam: PF00105; z1-c4; 1.	
DR	PIFNTS: P000647; SteroidHINER	
DE	PIFNTS: P00064; SteroidHINER	
DR	PROSITE: PS00031; NUCLEAR_RECEPTOR_1	
PW	Proceptor: Transcription regulation; DNA-binding; Nuclear protein;	
KW	Zinc-finger; Steroid-binding.	
FT	DOMAIN 15 170	MUTATING
FT	UNA_EIRC 171 236	MUT_FAP PEPTIDE-TYPE
FT	ZN_EIRC 171 191	CA-TYPE
FT	ZN_EIRC 207 231	CA-TYPE
FT	DE_EIRC 237 274	STERIOD-BINDING
SC	SPROUNCD 573 AA: 63420 MW: 92600	CDGCTIAGAGAGAGAGG CTR664:

```

Query Match: 55.28; Score 1549; DB 1; Length 573;
Post Local Similarity: 57.7%; Prod. No. 4.1e-116;
Matches: 319; Coverage: 62; Mismatches: 121; Indels: 46; Gaps: 10

QY 1 PSSYNQSSTPLPHNATVTPSSVYSHNEHYPMATYSEFVAMYSTIFSNVTEGCG--- 69
DB 33 PTMNTGALPALSMSHWAVCLSPPTSDSDYALATYSPILSH-----GGRAVP 82
QY 70 -----GGRAVSPNVWVW-----PCHSPVWVHQI SHVYRQKSNVCKASLEHTLVNKE 122
DB 83 ESPAPQSTISLSEFWMHGCHGVSPYALPHQOPIVYREPAHNSWAPKYLEH-----GQ 137
QY 123 ELKRVYNSNGVASTVGRSS-----KRAHRCVAVSDVAGSYGVMSGEGKAEFR 175
DB 138 AGTISKIAKKKAAEBSRSTSSVGGCFARKGDMHRCVAVGHVYASGVHVMSSGQKAFPR 197
QY 176 STYHNAYCTA.NC?..LHNPKNVYACTYECGYVWAPKCTSSPRCTGVYVRQSA 235
DB 198 SLCHNGVIGATVQCTLTKNRKSCQACRLPYCYEOMKMGQVPRFCYVCAHAFMR 257
QY 236 D-GLLHCAKCA---KSSGAPARVRELLIDALSPEDVILLAEVRPHV-LISRSAPR 289
DB 258 HIRPLAGTGGARPYGVRRVAVVPRDTGQKQNSALTTPROLINITAEPRPYLMEKELKPR 317
QY 349 TTASMMMLTEKLAKELEVMHSMARKTPGVVEISIPGVVPTESCMWAEVIMMGLMWRST 349
DB 318 TEDSMMSMISLADKEFLVLMISMAKKIPGVFEELDSJQVLEHLECCMEVLEMLGMSRSD 377
QY 450 HPGKILFAPVDVLDVDEKCEVEGLEIFDMILATTSRPRELLDKEVEICAKAMILNSS 409
DB 478 HPGKILFVLDKILNDEDSVCEVLEIFEDVLAATSPPRPIKIDREEVVLCIKALILPNP 437
QY 410 MYPLVTAI-VQADUSKRLAHLINAVVAVLVVAKSGISSGGCMRLANLIMLSHVSHA 468

```

RESULT	11			
HERL_CARAU				
ID	PREL_CARAU	STANDARD:	PRT:	FAA
AC	Q9W669;			
PT	01-OCT-2000 (Prl. 40, Created)			
DT	01-OCT-2000 (Prl. 40, Last sequence update)			
DT	01-OCT-2000 (Prl. 40, Last annotation update)			
DE	ESTROGEN RECEPTOR BETA-1 (ER-BETA-1).			
OS	Caassus auratus (Goldfish).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;			
OC	Cyprinodontes; Cyprinidae; Cyprininae; Cyprini;			
OX	NCBI_TaxID=7957;			

CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation.
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation.
CC	the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (Contact: license@ebi.ac.uk or send an email to license@sb.slb.ch).
CC	
DR	EMBL; AF061269; AAD26921.1; -
DR	HSSP; P03372; IHCP.
DR	InterPro; IPRO00324; -
DR	InterPro; IPRO01628; -
DR	PRINTS; PR00047; SFROIIDFNGR.
DP	PRINTS; PR00350; VITAMINR.
DR	PROSITE; PS00031; NUCLEAR_RECEPTOR_1.
KW	Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW	Zinc-finger; Steroid-binding.
FT	DOMAIN 12 .. 169 MNUATING.
FT	DNA_BIND 170 .. 245 NUCLEAR_RECEPTOR-TYPE.
FT	ZN_FING 170 .. 190 CA-TYPE.
FT	ZN_FING 206 .. 230 CA-TYPE.
FT	DOMAIN 236 .. 568 STEROID-BINDING.
SQ	SEQUENCE 568 AA; 6.3539 KM; ZDV/RD07/GNR/FY9 CRR64;

	Query Match	52.3%	Score 1466	M 1	Length 568	
	Best Local Similarly	56.0%	Fold. No. 17-109			
	Matches 290; Conservative	75;	Mismatches 127	Indels 26	Gaps 10	
Q7	SPSYNFQSILPLEHCSTYSSVVAHTRTAMFYAVWRYEELSHVNLPGPG-					70
	:					

BA Hawkins M.F., Thornton T.W., Crews D., Skipper J.K., Dotte A.,
 KA Thomas P.:
 "Identification of a third distinct estrogen receptor and
 KI reclassification of estrogen receptors in teleosts."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:10751-10756(2000).
 CC -1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
 CC ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
 CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
 CC SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 CC ALPHA (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN THE LIVER AND TESTES. LESS
 CC ABUNDANT IN THE OVARY AND BARELY DETECTABLE IN THE MUSCLE.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR3 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF298181; AAC16711.1;
 CC PROSITE: PS00031; NUCLEAR_RECEPTOR; 1;
 CC KEGG: K04401; Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger; Steroid-binding; MODULATING;
 CC TERNARY 1 161
 CC DNA_BIND 182 247
 CC ZN_BIND 182 202
 CC ZN_BIND 219 242
 CC ZN_BIND 248 653
 CC STERIOD-BINDING;
 CC SEQUENCE 673 AA; 74680 MW; 28394515757E3D01 CRC64;

Query Match 51.08; Score 1431; DB 1; Length 673;
 Best Local Similarity 53.58; Pred. No. 1.3e-106;
 Matches 103; Conservative 76; Mismatches 125; Indels 62; Gaps 18;

6 SPSSLSNRYVNMSSQSLPLEIKSIYDSSVDSHNEY-----PAMTYSPAVNYSIPSN 61
 DB SPGLI---PAVYS---PFLQMSHIVICPSPTTSSSHETYNHSHPLTFYSVSLSTYRPP- 84
 Q1 VTRLEQSTQCP-ATTSNVIWPTQTH--ISPIVWHPQSHLVAPP-KSQWTPAP----- 111
 DB ITN---SPSSIDPSLSAPFMSHPNHTLTPESIVVNDSPHAPLESKAHSTN 141
 Q12 SLPHHTLVNHTKKRVSSQ-----NKCASPIVIGDSKSHAPFCANVSAGSYAHYGV 164
 DB 142 ASSESTTEGNSLVKPSHVEYEDNMSSLDSSAV-----QADMHFCAVCHYASGYHGV 197
 Q13 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224
 DB 198 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 257
 Q14 CYRIKVRGKS-----ADEGLHAGKAKKSGCH-----AP--FVRELLTALASP 265
 DB 258 SYPAAPHPPTGPPPTGPPPTGPPPTGPPPTGPPPTGPPPTGPPPTGPPPTGPPPTGPP 317
 Q15 FAVVITTEFAVDVY-11SPSPAPFTFASMMSTTKLAFKRTVMISWAKKIDCEVELSL 324
 DB 418 FEFISRIHDAITPELYLMDLAKFTFASMMSTLNALFELVMSIWMKKIDCEVELSL 377
 Q16 FLYVLTLESVWMEVLMGLMSESTTHPKLLEAPVIFPREGCVAGSILEPMLATT 384
 DB 378 ADQINILKQCMELIMGLMMSVDRKLLFSPDKFLNREOCVAGSILEPMLAGT 437
 Q17 SFFRELKQKHYGVGVNMMILNNSMPLVYAI-QVADSSKRLAHLNANVTALVAVYAK 443
 DB 438 SFFRELKQKHYGVGVNMMILNNSMPLVYAI-QVADSSKRLAHLNANVTALVAVYAK 477

Q1 444 SGLSSQOOSMLANLMLLSHVRASR-32EALINE-YTAVVYV-11PRLMAHLYRGT 503
 DB 498 MGLTQOQOTIRLRLHLMILSHIHKVSKRMHLSHMEPRVVI-VIELLEMLAVNSQ- 556
 Q1 504 KSSITVSESPSPADRSK-KRGSNTPQ 528
 DB 557 -----GSPSSP 577

RESULT 14

ESR2_MACMU STANDARD; PRT; 279 AA.
 ID ESR2_MACMU
 AC Q9TTE5;
 DT 01-OCT-2000 (Poi 40, created)
 DT 01-OCT-2000 (Rel. 40, last sequence update)
 DT 01-OCT-2000 (Rel. 40, last annotation update)
 DE ESTROGEN RECEPTOR BETA (ER-BETA) (FRAGMENT).
 GN ESR2_OP NP342
 OS Macaca mulatta (Rhesus macaque).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 CX NCB1_TaxID:9544;
 CX 111
 RP SOURCE: FROM N.A.
 RX MEDLINE:20113291; PubMed:10645527;
 RA WO W.X., Ma X.H., Smith G.C.S., Nathanielsz P.W.;
 RT "Differential distribution of Eralpita and Eralpita mRNA in intracellular
 RT tissues of the pregnant rhesus monkey."
 RL Am. J. Physiol. 278:C190-C194(2000).
 CC -1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
 CC ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
 CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
 CC SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 CC ALPHA (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR3 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF119229; A05406.1;
 CC HSP: P03722; IERE.
 CC InterPro: IPR000546;
 CC InterPro: IPR01723;
 CC InterPro: IPR01728;
 CC Pfam: PF00104; hormone-rec; 1.
 CC PRINTS: PR00398; STEROIDHORMR.
 CC PRINTS: PR00546; TYRODIDHORMR.
 CC PROSITE: PS00031; NUCLEAR_RECEPTOR; PARTIAL.
 CC RECEPTOR: Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger; Steroid-binding; Phosphorylation;
 CC NON_TER 1
 CC DOMAIN 1
 CC MOD_RES 251
 CC NON_TER 279
 CC SEQUENCE 279 AA; 31105 MW; 858D9B7D01D0A301 CRC64;

Query Match 49.8%; Score 1398; DB 1; Length 279;
 Best Local Similarity 98.28; Pred. No. 1.8e-104;
 Matches 274; Conservative 32; Mismatches 22; Indels 0; Gaps 0;

Q1 239 QLVACGAVFSSCHADPVELLITPAI SPQVITITAFPIVLLISQSNAPFTASMMMS 297

Fri Apr 20 09:00:54 2001

us-08-906-365-2.rsp

Page 14

22	1233.6	73.2	2568	81	AC1583	AB1583 Sequence 1
23	1233.6	73.2	2568	81	AB075920	AB075920 Sequence
24	1167.4	69.2	1427	10	AF124790	AF124790 Homo sapi
25	1120.4	66.5	1621	11	AF042058	AF042058 Rattus no
26	1109	65.8	2152	11	MM081451	MM081451 Mus muscu
27	1099.6	65.2	1458	11	AX022460	AX022460 Sequence
28	1099.6	65.2	1458	11	MA057220	MA057220 Mus muscu
29	1099.6	65.2	1458	11	AB1588	AB1588 Sequence 6
30	1099.6	65.2	1458	11	AB075922	AB075922 Sequence
31	1056.4	62.7	1512	11	AF042059	AF042059 Rattus no
32	1056.4	62.5	1512	11	AB012721	AB012721 Rattus no
33	1053.2	62.5	1512	11	AF042060	AF042060 Rattus no
34	898.8	51.3	1504	11	AF042060	AF042060 Rattus no
35	879.4	52.2	3163	4	AF13513	AF13513 Sturinus v
36	865	51.3	3235	10	AF074593	AF074593 Homo sapi
37	834.8	49.5	1558	11	AF042061	AF042061 Rattus no
38	823.4	48.8	3539	4	AB036415	AB036415 Gallus ga
39	811.2	48.1	840	10	AF119229	AF119229 Macaca mu
40	764.8	45.4	1419	4	AF045149	AF045149 Corvus cor
41	519.2	30.8	3061	4	AB003356	AB003356 Anas plat
42	481.2	28.7	2293	4	AF041264	AF041264 Corvus cor
43	469.6	27.9	3508	4	CMY289883	CMY289883 Oncorhyn
44	453.4	26.9	127705	84	CM051011	CM051011 Homo sapi
45	453.4	26.9	178257	48	CNS01PH1	AL161756 Human chr

ALIGNMENTS

RESULT 1	AB006590	1740 bp	mRNA	PRI	05-FEB-1999
LOCUS	AB006590	1740 bp	mRNA	PRI	05-FEB-1999
DEFINITION	Homo sapiens mRNA for estrogen receptor beta, complete cds.				
ACCESSION	AB006590.1	CI:2911151			
VERSION	AB006590.1	CI:2911151			
KEYWORDS	estrogen receptor beta.				
SOURCE	Homo sapiens cDNA to mRNA.				
ORGANISM	Homo sapiens				
REFERENCE	Bouhassira, M.; Chodura, V.; Vertebata; Mammalia; Eutheria;				
AUTHORS	1 (bases 1 to 1740)				
JOURNAL	Submitted (13-APR 1997) to the EMBL/GenBank databases. Sumito				
REFERENCE	ogawa, Saitama Medical School, Department of 2nd Biochemistry; 38				
AUTHORS	Moriyama, Moriyama, Saitama Med. Sch., Saitama 350-0495, Japan				
JOURNAL	(E-mail: suogawa@saitama-med.ac.jp, Tel:81-492-76-1490,				
REFERENCE	2 (bases 1 to 1740)				
AUTHORS	ogawa, S.; Inoue, S.; Watanabe, T.; Hiroi, H.; Orimo, A.; Hosoi, T.;				
JOURNAL	Ouchi, Y. and Moriyama, M.				
REFERENCE	The complete primary structure of human estrogen receptor beta (HER				
AUTHORS	beta) and its heterodimerization with ER alpha in vivo and in vitro				
JOURNAL	Biochem. Biophys. Res. Commun. 243 (1), 122-126 (1998)				
REFERENCE	1 (bases 1 to 1740)				
AUTHORS	Local/Qualifiers				
JOURNAL	1. 1740				
REFERENCE	/organism="Homo sapiens"				
AUTHORS	/taxid="9606"				
JOURNAL	/product="estrogen receptor beta"				
REFERENCE	/protein_id="AA04953.1"				
AUTHORS	/db_xref="GI:2911152"				
JOURNAL	/translation="MDIKNSPSSLNSSSYNCSQSLPLEHSGYIPSSVDSHHEYP				
REFERENCE	AMTFEPAVMNYSIPSNVNLGGPGRTSPNVLPPIGHLSPVLRDVSHTIYAP				
AUTHORS	OKSPWEAKSLHTLTKRKVSGNRCAPVIGSGSKDAHFCVARSIVAGYGH				
JOURNAL	YGVMSDEGKAFKRSIQHNDYICDPATNOCTDKNRKSCQACCTPCYGVAVKCS				
REFERENCE	SRREKGVTLVKQSADEPLICACAKRSCGHARVELLIALSPQALILLKAE				
AUTHORS	PPHVLISRPAPTEASMMSTIKTADLELYHMSAKITGFVEISLFDVRLLEAC				
JOURNAL	WVEVLMKIMMSIDHPSKLPATDVLDRDCKGCEVIEIEMLLATVTSRFLKL				
REFERENCE	OKRETCYKAMLLNSMPLATVADDSKPKAHLNATVDALVWIASGTSISQ				
AUTHORS	OSRRLANLMLLSHVHASNKGEMHLLMMCKKNVYPVLLLEMLNAHLGKSSLT				

Query Match	99.98%	Score 1644.4	DB 8:	Length 1740:	
Best Local Similarity	99.98%	Pred. No. 0:			
Matches 1685:	Conservative	0: Mismatches	1: Indels	0: Gaps	0:
HASH COUNT	425 a	463 c	459 g	393 t	
ORIGIN	GSRCPAPUSKSKPSTONQSG				
Query	1	cagccattatactgcccacgaatcttggacacattataatgaccttattgacctct	60		
DB	6	CAGCCATTATACCTGGCCACGATCTTTGAGACATTATATGACCTTGTGGCTTCT	65		
Query	61	tacgaagatggtttctgagctgctatctgagtgagctgagctgagctgagctg	120		
DB	66	TGTAAGGTGCTTTTCTTACCTGCTTATCTCAGAGATGAGATATAAAACCTACCTAGC	125		
Query	121	cttaattct	180		
DB	126	CTTAATTCT	185		
Query	181	atatactacct	240		
DB	186	ATATACATACCT	245		
Query	241	agccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct	300		
DB	246	AGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	305		
Query	301	ggtcggagac	360		
DB	306	GPTGAGAC	365		
Query	361	gtatccatccatccatccatccatccatccatccatccatccatccatccatcc	420		
DB	366	GTSATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT	425		
Query	421	gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	480		
DB	426	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	485		
Query	481	ggaagacgttggc	540		
DB	486	GGAAGACGTTGGCG	545		
Query	541	gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	600		
DB	546	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	605		
Query	601	aagggccttctttaaaggaagcattcaaggaacattatattatattatattat	660		
DB	606	AAGGCTTTTAAAGGAAGCATTCAAGGACATATATATATATATATATATATAT	665		
Query	661	cagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	720		
DB	666	CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	725		
Query	721	tacgaagtggaagtggaagtggaagtggaagtggaagtggaagtggaagtgga	780		
DB	726	TACGAAGTGGAAGTGGAAGTGGAAGTGGAAGTGGAAGTGGAAGTGGAAGTGGA	785		
Query	781	cggagatgaggaatgaggaatgaggaatgaggaatgaggaatgaggaatgagga	840		
DB	786	CGGAGATGAGGAATGAGGAATGAGGAATGAGGAATGAGGAATGAGGAATGAGGA	845		
Query	841	gagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	900		
DB	846	GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	905		
Query	901	ctgagcctgctgagcctgagcctgagcctgagcctgagcctgagcctgagcct	960		
DB	906	CTGAGCCTGCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCT	965		


```

Db 867 CTGCTGTCAGCGATTACGATCCGATATCACTATGAGAGCTGCTGCTGTAAGAGATGA 926
G 602 GAGCTCTCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
Db 927 AGGCTCTCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 986
G 662 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
Db 987 AGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1046
G 722 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
Db 1047 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1106
G 782 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 841
Db 1107 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1166
G 842 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
Db 1167 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
G 902 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
Db 1227 TCAATCTCTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286
G 962 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
Db 1287 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1346
G 1022 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
Db 1347 TCAATCTCTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1406
G 1082 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1141
Db 1407 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466
G 1142 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201
Db 1467 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1526
G 1202 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1261
Db 1527 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
G 1262 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1321
Db 1587 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1646
G 1322 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381
Db 1647 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1706
G 1382 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441
Db 1707 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1766
G 1442 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1501
Db 1767 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1826
G 1502 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
Db 1827 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1886
G 1562 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1621
Db 1887 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1946
G 1622 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1681
Db 1947 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2006

```

```

G 1682 agtga 1686
Db 2007 ACTGA 2011

RESULT 3
LOCUS A65127
DEFINITION Sequence 24 from patent EP0798378.
ACCESSION A65127
VERSION A65127.1 GI:4530991
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1898)
AUTHORS Mosseiman, S. and Dijkema, K.
TITLE Estrogen receptor
JOURNAL Patent, EP 0798378-A, 01-05-1997;
AKZO NOBEL, NY (NL)
COMMENT other publication: CA 220423, 19970926,
FEATURES location/Qualifiers
source 1..1898
/organism="unidentified"
/db_xref="taxon:32044"

BASE COUNT 452 a 506 c 501 g 439 t
ORIGIN

Query Match 99.9% Score 1667.4 db:gi Length 1898;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1668; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

G 18 caggaatcttggagacatataatgaacatttgacattcttgaagttgtttctga 77
Db 1 CAGCAATCTTTGACACATTAATATGACCTTTGCTCTTCTTGGAGAGCTTTCTGA 60
G 78 gctctctatctgaagatggatataaagaatcagcctctgacttaattctctctc 137
Db 61 GCTCTATCTCAAGACATCATATMAAATAACATGACCATTAATCTTCTCTCTG 120
G 138 ctatgaatctgaagatctgaagatctgaagatctgaagatctgaagatctgaag 197
Db 121 CTATGAATCTGAATCTTATCTTACCTCTGAGCAAGCTCTATATACATACCTTCT 180
G 198 ctatctatgaagacagacatataatgaacatgaagatctgaagatctgaagat 257
Db 181 CTATGAGACAGCAGCATGATATCCAGCATGACATTCATAGCCCTCTCTATCA 240
G 258 ttacagatcttggagacatataatgaacatttgacattcttgaagttgtttctga 317
Db 241 TTACAGCATTTCCAGCATTTCTGACATCTGCAAGAGAGAGAGAGAGAGAGAGAG 300
G 318 ccccaatgtttatgagacagacatataatgaacatttgacattcttgaagttgt 377
Db 301 CCCAATGTTCTTGTGTGCAACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
G 378 atcagatctgaatgagacatataatgaacatttgacattcttgaagttgtttctga 437
Db 361 ATCAGATCTCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
G 438 cagcttacctgttaagagagagagatataatgaacatttgacattcttgaagttgt 497
Db 421 CAGCTTACCTTTAAATACAGACAGATATTAAGAGAGAGAGAGAGAGAGAGAGAG 480
G 498 cctcttactgttgcagagatctgaagatctgaagatctgaagatctgaagatctga 557
Db 481 CCTCTTACTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
G 558 cagatctgaatctgaatctgaagatctgaagatctgaagatctgaagatctgaag 617

```


[illegible]

DEFINITION	Human sapiens estrogen receptor beta2 splice variant mRNA, complete cds.
ACCESSION	AF051428
VERSION	AF051428.1
KEYWORDS	GI.2967558
SOURCE	human.
ORGANISM	human sapiens Eukaryota; Metazoa; Chordata; Grafiata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 2041) Moore,J.T., McKee,D.D., Moore,L.B., Jones,S.A., Su,J.-L., Horne,E.L., Kilewer,S.A., Coleman,J.M. and Willison,T.M. Cloning, ligand binding and functional Activity of human Estrogen Receptor beta Isoforms Unpublished
JOURNAL	2 (bases 1 to 2041)
AUTHORS	Moore,J.T., McKee,D.D., Moore,L.B., Jones,S.A., Su,J.-L., Horne,E.L., Kilewer,S.A., Lehmann,J.M. and Willison,T.M. Direct Submission Submitted (25 FEB 1998) Molecular Sciences, Mayo Wellcome, 5 Mayo Drive, RTP, NC 27709, USA Location/Qualifiers 1..2041
FEATURES	source CDS /organism="Homo sapiens" /db_xref="taxon:9606" /translation-type="testis" 419..1906 /codon_start=1 /product="estrogen receptor beta2 splice variant" /protein_id="AAC05751.1" /db_xref="gi:2961559"
Query Match:	88.8%; Score 1496.4; Db 10; Length 2041;
Best Local Similarity:	93.9%; Freq: No. 0;
Matches 1497:	Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Df	2 agccattatctgcccagcaatctttagacatatagaactttgtgcctctc 61
327	accgatattactggccagcatcttgaggacatttatatgcccttgtgccttcct 86
OY	62 gccaagtgtttctcagctgatatctcgaagatgaaatctaatctaacatctaac 121
Db	387 gcagcgctgtttctcagcgttatcttcacaaagatgataaaaaatgcacatctaac 446
OY	122 ttaatttctctctctcctaacaactgaagatgatatcttacctggaagacagctcca 181
Db	447 tttaatttctctctctctctacacagctgtagatattgcatctttaccctctgaaacnaggctcca 506
OY	185 tatcatactcctctcctatatgtagcagccagcaatgatatctccagctctctata 241
Db	507 tatacatatgcctctcctcctctctatgtatagatgacacacacatgaaatccacatctctata 566
OY	242 gccctgctgtgttgaattacagatctctcgaagatgtcaetaaetttgaagaatgagctga 401
Db	567 gccctgcctgtcgaatgaaattacacacattccacagatattctactaaccttgaaaatgagctga 626
OY	302 gtctgacacacacagcgaatgtgtgtgacgaagcctggagagctctctctctttag 461
Db	627 gtctgacacacacagcgaatgtgtgtgtgacgaagcctggagagctctctctctttag 686
OY	362 tggctcacctgcagattatcactctgtatgqqaacactcgaatgagctctctgtgtgag 421

[illegible][illegible]

[illegible][illegible]

```

UY 587 cgttgttgaagatgtaagggcttllttaaagaagatccaaggaatgattatatt 646
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 cgttgttgaagatgtaagggcttllttaaagaagatccaaggaatgattatatt 420
UY 647 gttcaaatcaaatgattgataatcaataaagaaggaagatgccagggctggc 706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 gttcaaatcaaatgattgataatcaataaagaaggaagatgccagggctggc 480
UY 707 gaattcgaagatgtaaggaatgtaaggaatgtaaggaatgtaaggaatgta 766
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 gaattcgaagatgtaaggaatgtaaggaatgtaaggaatgtaaggaatgta 540
UY 767 gataagggcttgaaggaagagaagatgataaggaaggaagatgataagga 826
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 gataagggcttgaaggaagagaagatgataaggaaggaagatgataagga 600
UY 827 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 660
UY 887 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 946
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 720
UY 947 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 780
UY 1007 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1066
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 840
UY 1067 tgaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 tgaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 900
UY 1127 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 960
UY 1187 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1020
UY 1247 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1080
UY 1307 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1140
UY 1407 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1407 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1260
UY 1427 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1427 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1280
UY 1447 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1506
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1447 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1320
UY 1547 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1547 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1380
UY 1607 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1666
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1607 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 1440

```

```

UY 1667 agaaccacagctcagtgga 1686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1441 agaaccacagctcagtgga 1460

RESULT 11
AF075921 1460 bp DNA PAT 30-AUG-2000
LOCUS AF075921 4 from patent US 5958710.
DEFINITION Orphan receptor
ACCESSION AF075921
VERSION AF075921.1 GI:10002667
KEYWORDS
SOURCE
  ORGANISM Unknown.
  UNCLASSIFIED.
  REFERENCE 1 (bases 1 to 1460)
  AUTHORS Kuiper, G., Denmark, E.L.K. and Gustafsson, J.
  TITLE Orphan receptor
  JOURNAL Patent: US 5958710-A 4 28-SEP-1999;
  FEATURES
    source
      1..1460
      /organism="Unknown"
BASE COUNT 352 a 379 c 411 g 318 t
ORIGIN

Query Match 86.2% Score 1453.6; DB H1; Length 1460;
Best Local Similarity 99.7%; Frq. No. 9;
Matches 1456; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

UY 227 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 286
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 60
UY 287 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 346
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 120
UY 347 acccttcccttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 406
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 acccttcccttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 180
UY 407 gttccttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 gttccttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 240
UY 467 aaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 aaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 300
UY 527 atgctcacttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 atgctcacttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 360
UY 587 cgttgttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 646
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 cgttgttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 420
UY 647 gttcagcttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 gttcagcttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 480
UY 707 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 766
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 gatttgaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 540
UY 767 gataagggcttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 826
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 gataagggcttgaaggaaggaagatgtaaggaaggaagatgtaaggaag 600
UY 827 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 ccaaggaagatgtaaggaaggaagatgtaaggaaggaagatgtaaggaag 660

```


[illegible]

KEYWORDS
SOURCE
ORGANISM
PREFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
unidentified.
unidentified
unclassified.
1 (bases 1 to 1434)
Mosseelman, S. and Dijkema, R.
Estrogen receptor
Patent: EP 0/98378-A 01-0675-1997;
AKZO NOBEL NV (NL).
other publication CA 2706143 1/25/97
other publication AU 1652197 19971002.
Location/Qualifiers
1..1434
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 347 a 343 c 406 g 308 t
ORIGIN

[illegible]

Fri Apr 20 09:00:44 2001

us-08-906-365-1.rge

Page 18

PT human oestrogen receptor-beta interactive compounds
XX
PS Claim 1; Page 41-42; 49pp; English.
82

This cDNA clone encodes a full-length human oestrogen receptor-beta (hER-beta, see W9756) containing 45 previously unknown N-terminal amino acid residues that are believed to contribute to the transcription activation function of the receptor. Human testis cDNA was subjected to PCR using primers (see X2436-66) designed to amplify the previously known partial human sequence with 2, missing N-terminal residues (Met and Thr) found in rat ER-beta, and to supply an artificial Kozak translational initiation sequence. A fragment of approximately 1500 bp was obtained, and was cloned into vector pCDNA3 for sequencing. The 5' sequence was confirmed by PCR amplification from a human ovary cDNA library and by 5'RACE (see X2437-71). ER-beta is selectively expressed in the thymus, spleen, ovary and testes. The invention encompasses hER-beta polynucleotides and polypeptides, particularly peptides which include residues 1-45 of hER-beta. The invention also provides expression systems in which transcriptionally active hER-beta or fragments can be produced, as well as screening methods for identifying hER-beta agonists and antagonists (including tissue-specific oestrogens and anti-oestrogens) and hER-beta co-activators and inhibitors.

Sequence: 1686 BP, 413 A, 449 C, 441 G, 383 T, 0 other;

Query Match	100.0%	Score	1686	IP	20	Length	1686	
Post-local Similarity	100.0%	Pred. No.	0					
Matches 1686	Conservative	0	Mismatches	0	Indels	0	Gaps	0

QY	1	ccccccattatcttgcgcacgaagatctctttggaacattatataatgaccttttggctctctc	60
Db	1	ccccccattatcttgcgcacgaagatctctttggaacattatataatgaccttttggctctctc	60
QY	61	taaaagagattttctcaactgctatctcaagaacatgatatataaaaatcaccatctagc	120
Db	61	tcaagaggttttctcagctgctatcttcaagaactgatatataaaaatcaccatctagc	120
QY	121	cttaattctctctctctctctcaactgcagtgatctaatcttaacccctggagcaagctcc	180
Db	121	cttaattctctctctctctctcaactgcagtgatctaatcttaacccctggagcaagctcc	180
QY	181	ataataacacctctctctctctatctgataacgcacacatgatatccagccatgaaattcat	240
Db	181	ataataacacctctctctctctatctgataacgcacacatgatatccagccatgaaattcat	240
QY	241	agccctgcctgataataatatacagactctccagcaatgtcacataacttgaagaatggacct	300
Db	241	agccctgcctgataataatatacagactctccagcaatgtcacataacttgaagaatggacct	300
QY	301	ggttcgtgcagacacacacagccccaattttgttttgggcacacactgggcacctttcccttta	360
Db	301	ggttcgtgcagacacacacagccccaattttgttttgggcacacactgggcacctttcccttta	360
QY	361	gtatgtccac	420
Db	361	gtatgtccac	420
QY	421	gcaaaatcagctcaaac	480
Db	421	gcaaaatcagctcaaac	480
QY	481	ggaagacacgtctgcgcagac	540
Db	481	ggaagacacgtctgcgcagac	540
QY	541	gctgctctgcagagatataac	600
Db	541	gctgctctgcagagatataac	600
QY	601	aaagcctcttttataaagaagacataaagacataatgatatatttgcagctaccacaaat	660

[illegible]

PK	08-MAY-1996:	96GB-0009576.
PK	08-SEP-1995:	96GB-0018272.
PK	15-MAR-1996:	96GB-0005550.
PK	11-APR-1996:	96GB-0007512.
PA	(KAF0-) KAF0 R10 AB.	
XX		
XX	Enmark E, Gustafsson J, Kuiper GG:	
XX	WPJ: 1997-192842/17.	
XX	P-PSDB: W14724.	
PT	New isolated oestrogen receptor beta used to develop probes for	
PT	treatment of gromphs, CNS diseases, osteoporosis or cardiovascular	
PT	disease	
XX		
PS	Claim 6: Fig 13b; 45pp; English.	
XX		
CC	A cDNA clone (T67843) codes for a novel human oestrogen receptor	
CC	related nuclear receptor (W17524), designated ER-beta. It was	
CC	isolated from a human testis cDNA library using rat ER-beta cDNA	
CC	(see 3150 76:842) as probe. The human ER-beta gene was mapped	
CC	to chromosome 14 region 14q22-23. Rat, human and mouse ER-beta	
CC	(W14723-25) can be used to isolate molecules for use in the	
CC	treatment of cardiovascular diseases, central nervous system	
CC	diseases, osteoporosis, prostate or ovarian cancer or benign	
CC	prostatic hyperplasia and to test environmental chemicals for	
CC	oestrogenic activity.	
XX		
SO	Sequence 1460 bp; 352 A; 479 C; 411 G; 418 T; 0 other.	
Query Match		
Best Local Similarity 86.2% Score 145.7 Dip 181 Length 1460		
Matches 1456; Conservative 0; Mismatches 4; Indels 0 Gaps 0		
QY	227 ccatgacatctctatgacctctatctatgaatgaataagatctccaaatgctgaact	289
DB	1 cctagacatctctacacctctgctgctgaatgaataagatctccaaatgctgaact	60
QY	267 tggagagtgagatctgctctggcagacacaaacaaatctcttctggcgaactggac	446
DB	61 tggagagtgagatctgctctggcagacacaaacaaatctcttctggcgaactggac	120
QY	347 acctctctctcttactggtctatgctgcttatacaatctatcttgcgaacctcaaaa	406
DB	121 acctctctctcttactggtctatgctgcttatacaatctatcttgcgaacctcaaaa	186
QY	407 gtccctgtgtgaagaaagatcgtctgaagacacgtcttaccgttataagcaaaagatga	466
DB	141 gtccctgtgtgtgaagaaagatcgtctgaagacacgtcttaccgttataagcaaaagatga	240
QY	467 aaagaaagattttatgaagacacgttctgcgcacacgtcttactatgctcctttcaaaagaa	527
DB	241 aaagaaagattttatgaagacacgttctgcgcacacgtcttactatgctcctttcaaaagaa	600
QY	527 atgctcaatctctggatgtatgaaagatgaatctgaatctgaatctgaatctgaat	589
DB	301 atgctcaatctctggatgtatgaaagatgaatctgaatctgaatctgaatctgaat	600
QY	587 cctatgaagagatgttaagaccttttttaaaagaaagatctcaaaatgaatgaatctatatt	646
DB	361 cctatgaagagatgttaagaccttttttaaaagaaagatctcaaaatgaatgaatctatatt	420
QY	647 gtcctgctatcaaatctgtctacaatctggtataaaatctgcgcacaaatctcaaaacctgac	706
DB	421 gtcctgctatcaaatctgtctacaatctggtataaaatctgcgcacaaatctcaaaacctgac	480
QY	707 gaacttcgaagatgttatacaatctggtatgaatctgaatctccgaagaaagatctgtg	766
DB	481 gaacttcgaagatgttatacaatctggtatgaatctgaatctccgaagaaagatctgtg	740
QY	767 gctacagcctctctgcgaagacgaagatgtgcgcgaagatctgctctgctctgcgcgaag	826

[illegible][illegible]

[illegible][illegible]

[illegible][illegible]

Sequence 2568 BP; 634 A; 654 C; 663 G; 617 T; 0 other;

Query Match	73.2%	Score 12345	DR 18	Length 2568
Best local Similarity	84.8%	Pred. No. 0		
Matches 1419	Conserved 0	Mismatches 249	Indels 6	Gaps 3

UY	16	caaaagaaactttaaagaacatataaagaaacttttgcgcctctcttgcagaagtgcttttc	75
DB	211	ccccgaagctctctggaataaataat--gttcaatctgaacctctctccaaagagcttttc	268
UY	76	caaatgaa--cttcaaaatataaataaactcaacatctaaacttaattctct	132
DB	269	caatctgagacctctcaaaagcttgaggttcaaaactcaacgttcgaagttagtttccct	327
UY	133	ctctctaaactctgaatcaatcctacttaacctggagagagctccataataacct	192
DB	328	gcctctctaaactgaatgaagcaatctctactccctctgagctccctccatctacacct	387
UY	193	tcctctctatgtaaaagagccatataaataaccagcttgacatctctaaacctgtgtg	252
DB	388	tcctctctatgtaaaagagccatataaataaccagcttgacatctctaaacctgtgtg	447
UY	253	atgaatttaagaacttccagaaatgctcaacttaacttaagaagtgaggtctgttcggagacc	312
DB	448	atgaatttaagaatgctcaagaacacacttaacctgagcttggagctgtccagctgagc	507
UY	313	atgaagcccaaatgctttgagccaaaccttgagcaacctctcccttaagtgtgtccctgcgc	372
DB	508	atgaagcccaaatgctttgagctaaacctcttgaggcaacctgtctcttttaagacccttgc	567
UY	373	caattatcaactcttaatgaggaacctctcaaaagagctccctctgtgtgaagaacaaatcgcta	432
DB	568	caactatctgtctctctatagcaaaacctcaaaagagctctctgtgtgtgaagaacaaatcgcta	627
UY	433	gaacacacactttaaactgttaaaataaagagacacatgtaaaagagagttatgtgaaccttgc	492
DB	628	gagacacactttaaactgttaaaataaagagacacatgtaaaagagagttatgtgaaccttgc	687
UY	493	gccaagcccttgaactgtgtcccaagcttcaaaagagagatgctcaactctgcgcgtctgcagc	552
DB	688	gccaagcccttgaactgtgtcccaagagagagatgctcaactctgcgcgcgtctgcagc	747
UY	553	gaattacgaatcaggaatactaatgagatctgcgtctgtgtgaagaatgttaagaccttttt	612
DB	748	gaattacgaatcaggaatactaatgagatctgcgtctgtgtgaagaatgttaagaccttttt	807
UY	613	ataaaagaacatctcaagagcatataatgattatatttgcacagctacaatactagttacaactc	672
DB	808	ataaaagaacatctcaagagcatataatgattatatttgcacacagcaatcaagttaccata	867
UY	673	gataaaagccgagcaaaagatctgcagagccctgcgcgaatttcggagaggttttaagaagtggga	732
DB	868	gataaaagccgagcaaaagatctgcagagccctgcgcgaatttcggagaggttttttaagaagtggga	927
UY	733	atgaatgaaggtatggtctcccaaaataaagaaatgtaggtctccgcttctgtgtgaacagagga	792
DB	928	atgaatgaaggtatggtctcccaaaataaagaaatgtaggtctccgcttctgtgtgaacagagga	987
UY	793	aatgcgcgcagagcagctgtgcactgtgcgcgcgagaaagccaaagatgtgcgcgcagcgcgc	852
DB	988	aatgcgcgcagagcagctgtgcactgtgcgcgcgagaaagccaaagaaatgtgtgtgtgtgtgtgtgt	1047
UY	853	cgaatctgagagagcagctgtgcactgt	912
DB	1048	cgaatctgagagagcagctgtgcactgt	1107
UY	913	gaagatgaagccgtcccatatgactgataatcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	972
DB	1108	gaagatgaagcccatatgactgataatcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1167
UY	973	atgaatgaagcccatatgactgataatcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1032

Dh	1168	atgagtagctgctccctaacaagctctggtggtgacgagaaatcgtttgacacatggttctggtcctggtc	1227
Qy	1033	aagaagaattcccgctcttttgtagaacctacgcctgtttgacccaattgcaggtctcttgagaagc	1092
Dh	1228	aaagaaataccctggtctgttgtagcctcagcctgttttgacccaattccggtctcttgagaagc	1287
Qy	1093	tatttgagatggagagtttaataatctggagctaaatctgaagctcaattgacccaacacgagcag	1152
Dh	1288	tgcctggatcgagagcgctcaatacagtgaggaactaaatctaaatctaaatctaaatctaaatct	1347
Qy	1153	cctacatcttgcctccagatcttctgtctgacacaggaatgaaggaataatctgtataagaatct	1212
Dh	1348	cctacatcttgcctccagacctcgtttctgacacaggaatgaaggaataatctgtataagaatct	1407
Qy	1213	cctggaacatttgacagctcctgaggaactctcgaagcttcctgaggaattaaactctga	1272
Dh	1408	cctggaacatttgagaatcctgaggaactctcgaagcttcctgaggaattaaactctga	1467
Qy	1273	cacaaagaatgactctgtgtaagaggaatgactcctgctcaattcctgaatgtagacctctga	1332
Dh	1468	cacaaagagtagctcctgtgtaagaggaatgactcctgctcaattcctgaatgtagacctctga	1527
Qy	1333	gtcaagaagaccctggatgctgacagagacgtcgaagactgctcacttgctaaacgtcgtg	1392
Dh	1528	gtctctgcaaacacagagaggaagaatgacctggaagactgacacacttaacacgagctg	1587
Qy	1393	accgagctgtcttggtcttggtgagatgtccaaagacgagatcctcctccgaagagaaactga	1452
Dh	1588	acaagatgcctggtctggtgagatgtccaaagacgagatcctcctccgaagagaaactga	1647
Qy	1453	cgcctcgctaacctccctcgaatgctcctgtcccaagctcaagcaatgcgaatgaacaaagcag	1512
Dh	1648	cgaactggacacctccctcgactctctctctccctgaagagcaatcagtaaacaaagcagctg	1707
Qy	1513	gaacatctgctcacaatcagaagctgcaaaaatgttgtccgaattatgaactgtctctcaga	1572
Dh	1708	gaacatctgctcacaatcagaagctgcaaaaatgttgtccgaattatgaactgtctctcaga	1767
Qy	1573	atgcttgattgcctcgctgctctgagagctgcaaatccctccatccacaggaatccaggtctcagc	1632
Dh	1768	atgcttgattgcctcacaagctctcgaagctgcaaatccctccatccacaggaatccaggtctcagc	1827
Qy	1633	cctgagacagacaaatgaagaacaaagaggtctccgaagcccaagctctcagtgag	1686
Dh	1828	ctcaacagagagacaaatgaagaacaaagaggtctccgaagcccaagctctcagtgag	1881
RESULT 11			
ID	X26685	standard; DNA; 2568 bp.	
XX	XX		
AC	X26685;		
XX	XX		
DT	17-JUN-1999	(first entry)	
XX	XX		
DE	DNA encoding an estrogen receptor-beta.		
XX	XX		
KW	differential estrogen receptor; ER-alpha; ER-beta; activation;		
XX	XX	API site; therapy: ss.	
OS	Rattus sp.		
XX	XX		
PH	Key	location/Qualifiers	
FT	CDS	424..1881	
XX	XX	/ftag a	
PN	W09911760-A1.		
XX	XX		
PD	11-MAR-1999.		
XX	XX		
PT	31-AUG-1998;	98MO-US18030.	
XX	XX		
RR	04-SEP-1997;	97US-0923708.	

[illegible][illegible]

```

RESULT 2
US-08-836-620A-1
Sequence 1, Application US/08836620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcaltto Rel-base #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08-836-620A
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/03933
FILING DATE:
APPLICATION NUMBER: GR 9618272.1
FILING DATE: 08-SEP-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GR 9605550.4
FILING DATE: 15-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9607532.0
FILING DATE: 11-APR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GR 9606576.5
FILING DATE: 08-MAY-1996
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2568 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
US-08-836-620A-1

Query Match 73.2%; Score 1293.6; E-2; Length 2568;
Best Local Similarity 84.8%; Pval 0;
Matches 1419; Conservative 0; Mismatches 249; Indels 6; Gaps
1;

QY 16 ccacgaatcttgaagaacatataatgaacttgaactcttcttgaagatattct 75
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 211 cccatgaagctctgcgaacacataat--gttcattctgtacacgtcttgcacaaagagatttct 268
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 76 ccactgacta--tctcaagacatgatatataaanaactcaacatctctatcttaattctctt 352
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


[illegible]

```

1  Sequence 7, Application US/08481970
2 Patent No. 5859310
3 GENERAL INFORMATION:
4 APPLICANT: Gossen, Manfred
5 APPLICANT: Bujard, Hermann
6 APPLICANT: Sallfeld, Jochen
7 APPLICANT: Voss, Jeffrey
8 TITLE OF INVENTION: Animal Transgenic for a Telomeric Line-Controlled
9 NUMBER OF SEQUENCES: 10
10 CORRESPONDENCE ADDRESS:
11 ADDRESSED: Lathe & Cockfield
12 STREET: 60 State Street, Suite 510
13 CITY: Boston
14 STATE: Massachusetts
15 COUNTRY: USA
16 ZIP: 02109-1875
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: ASCII text
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/481,970
24 FILING DATE:
25 CLASSIFICATION: 800
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: 08/250,452
28 FILING DATE: 14-JUN-1994
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: 08/076,327
31 FILING DATE: 14-JUN-1993
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Giulio A. DeConti, Jr.
34 REGISTRATION NUMBER: 31,503
35 REFERENCE/DOCKET NUMBER: BH1-01 30P2
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: (617) 227 7400
38 TELEFAX: (617) 227-5941
39 INFORMATION FOR SEQ ID NO: 9:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 4963 base pairs
42 TYPE: nucleic acid
43 STRANDEDNESS: double
44 TOPOLOGY: circular
45 MOLECULE TYPE: DNA (genomic)
46 ORIGINAI SOURCE:
47 ORGANISM: Human cytomegalovirus
48 IMMEDIATE SOURCE:
49 CLONE: pUHD BGH4
50 US-08-481-970-9
51
52 Query Match 23.2%, Score 390.9, E-Val 1e-06, Length 4963
53 Best Local Similarity 62.4%; Pred.No. 3,70-97;
54 Matches 553, Conservation % MisMatches 185 Indels 39 Gaps 4
55
56 5' aggttgctcaactctgagcgccttggcgatctttatcgaattacatttagcttc 582
57 ttttttttttttttttttttttttttttttttttttttttttttttttttttttt 582
58 tttaaaacacacctccacatgtagtgcacacacacacacacacacacacacacac 1062
59 cagtctgctgcatgagctgagctgagctgagctgagctgagctgagctgagctg 642
60 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 642
61 tggcttctgacacacacacacacacacacacacacacacacacacacacacac 1122
62 attcttctgacacacacacacacacacacacacacacacacacacacacacac 1182
63 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1182
64 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1242
65 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1242
66 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1302
67 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1302
68 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1362
69 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1362
70 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1422
71 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1422
72 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1482
73 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1482
74 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1542
75 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1542
76 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1602
77 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1602
78 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1662
79 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1662
80 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1722
81 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1722
82 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1782
83 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1782
84 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1842
85 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1842
86 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1902
87 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1902
88 atcttctgacacacacacacacacacacacacacacacacacacacacacac 1962
89 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 1962
90 atcttctgacacacacacacacacacacacacacacacacacacacacacac 2022
91 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 2022
92 atcttctgacacacacacacacacacacacacacacacacacacacacacac 2082
93 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 2082
94 atcttctgacacacacacacacacacacacacacacacacacacacacacac 2142
95 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 2142
96 atcttctgacacacacacacacacacacacacacacacacacacacacacac 2202
97 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 2202
98 atcttctgacacacacacacacacacacacacacacacacacacacacacac 2262
99 ttttttttttttttttttttttttttttttttttttttttttttttttttttt 2262
100 atcttctgacacacacacacacacacacacacacacacacacacacacacac 2322

```



```

940 CTATAGCTGAGCGCCGCAATGCTGATCTCTTGGCGCCGAGTGGTGGCTTCAAGGAG 946
1423 TTGGATGCTGGAGGCGGCGGATATCTGATGCGGAGATGAGATGCTGACGAGCCCTGAGTAA 1432
947 AGCTGAGGATGATGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1026
1434 GCTTGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1542
1027 TGGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1086
1543 TGGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1602
1087 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1146
1603 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1662
1147 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1206
1603 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1722
1207 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1266
1223 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1782
1267 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1326
1783 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1842
1327 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1383
1843 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1902
1884 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1962
1903 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1962
1444 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1503
1903 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1962
1504 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1563
2023 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 2082
1564 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1603
2083 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 2122

```

RESULT 8
 US-08-693-940-2
 Sequence 2: Application US/08693940
 Patent No. 6335027
 GENERAL INFORMATION:
 APPLICANT: Yee, Jinn-Kuan
 APPLICANT: Friedman, Theodore
 APPLICANT: Chen, Shin-Tai
 TITLE OF INVENTION: Inducible Expression System
 TITLE OF INVENTION: Useful in the Generation of Packaging Cell Lines for
 TITLE OF INVENTION: Pseudotyped Retroviral Vectors
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bozicevic & Reed, LLP
 STREET: 285 Hamilton Ave, Suite 200
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94301
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSO for Windows Version 2.0

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,940
FILING DATE: 07 Apr 1996
CLASSIFICATION: A35
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Francis, Carol L.
REGISTRATION NUMBER: 36,513
REFERENCE/EXEMPT NUMBER: 6510-055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1956 base pairs
TYPE: nucleic acid
STANDEDNESS: single
TOPOL/KEY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1..1953
OTHER INFORMATION:
US-08-693-940-2

```

Query Match 16.0%; Score 285.6; DB 3; Length 1956;
 Best Local Similarity 63.8%; Pred. No. 1,30-68;
 Matches 467; Conservative 0; Mismatches 259; Indels 6; Gaps 2;

```

878 CCGGAGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGAT 937
1084 CCGGAGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGAT 1153
938 TC AGCGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGAT 994
1154 CCGGAGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGAT 1213
995 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1054
1214 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1273
1055 AGCGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1114
1274 ATTGAGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1333
1115 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1174
1334 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1393
1175 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1234
1354 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1413
1235 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1294
1454 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1513
1295 AGCGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1353
1514 AATGAGGATGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1573
1354 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1411
1574 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1633
1412 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1471
1634 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1693
1472 TGGCGCGGAGCGAGGATGATGCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1531

```


Fri Apr 20 09:00:46 2001

us-08-906-365-1.rni

Page 13

Search completed: Apr 19, 2001, 05:16:20
Index time: 41.576 sec

•

•

•

GenScore version 4.5
Copyright (c) 1994 - 2000 CompuGen Ltd.

us nucleotide nucleotide search, using sw model

Run on: April 18, 2001, 09:44:24 : Search time 2082.96 Seconds

(without alignments)
5672.046 Million cell updates/sec

File: us-08-906-365-1

Perfect Score:

1686

Sequence: 1 caagcattatctatcttccac.....aaagccacagctctcagga 1686

Score for table:

IDENTITY.MP

Gapop 10.0 / Gapext 1.0

Searched: 7991742 seqs, 450474486 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum hit seq length: 0

Maximum hit seq length: 200000000

Post processed: Maximum March 98

Listing first 45 summaries

Database:

EST.*

1: qb_0811.*
2: qb_0812.*
3: qb_0813.*
4: qb_0814.*
5: qb_0815.*
6: qb_0816.*
7: qb_0817.*
8: qb_0818.*
9: qb_0819.*
10: qb_0810.*
11: qb_0811.*
12: qb_0812.*
13: qb_0813.*
14: qb_0814.*
15: qb_0815.*
16: qb_0816.*
17: qb_0817.*
18: qb_0818.*
19: qb_0819.*
20: qb_0820.*
21: qb_0821.*
22: qb_0822.*
23: qb_0823.*
24: qb_0824.*
25: qb_0825.*
26: qb_0826.*
27: qb_0827.*
28: qb_0828.*
29: qb_0829.*
30: qb_0830.*
31: qb_0831.*
32: qb_0832.*
33: qb_0833.*
34: qb_0834.*
35: qb_0835.*
36: qb_0836.*
37: qb_0837.*
38: qb_0838.*
39: qb_0839.*
40: qb_0840.*
41: cm_0810.*
42: cm_0811.*
43: cm_0810.*

44: cm_0810.*
45: cm_0811.*
46: cm_0812.*
47: cm_0813.*
48: cm_0814.*
49: cm_0815.*
50: cm_0816.*
51: cm_0817.*
52: cm_0818.*
53: cm_0819.*
54: cm_0820.*
55: cm_0821.*
56: cm_0822.*
57: cm_0823.*
58: cm_0824.*
59: cm_0825.*
60: cm_0826.*
61: cm_0827.*
62: cm_0828.*
63: cm_0829.*
64: cm_0830.*
65: cm_0831.*
66: cm_0832.*
67: cm_0833.*
68: cm_0834.*
69: cm_0835.*
70: cm_0836.*
71: cm_0837.*
72: cm_0838.*
73: cm_0839.*
74: cm_0840.*
75: cm_0841.*
76: cm_0842.*
77: cm_0843.*
78: cm_0844.*
79: cm_0845.*
80: cm_0846.*
81: cm_0847.*
82: cm_0848.*
83: cm_0849.*
84: cm_0850.*
85: cm_0851.*
86: cm_0852.*
87: cm_0853.*
88: cm_0854.*
89: cm_0855.*
90: cm_0856.*
91: cm_0857.*
92: cm_0858.*
93: cm_0859.*
94: cm_0860.*
95: cm_0861.*
96: cm_0862.*
97: cm_0863.*
98: cm_0864.*
99: cm_0865.*
100: cm_0866.*
101: cm_0867.*
102: cm_0868.*
103: cm_0869.*
104: cm_0870.*
105: cm_0871.*
106: cm_0872.*
107: cm_0873.*
108: cm_0874.*
109: cm_0875.*
110: cm_0876.*
111: cm_0877.*
112: cm_0878.*
113: cm_0879.*
114: cm_0880.*
115: cm_0881.*
116: cm_0882.*

117: em_estp16:*
 118: em_estp17:*
 119: em_estp18:*
 120: em_estp19:*
 121: em_estp20:*
 122: em_estp21:*
 123: em_estp22:*
 124: em_estp23:*
 125: em_estp24:*
 126: em_estp25:*
 127: em_estp26:*
 128: em_estp27:*
 129: em_estp28:*
 130: em_estp29:*
 131: em_estp30:*
 132: em_estp31:*
 133: em_estp32:*
 134: em_estp33:*
 135: em_estp34:*
 136: em_estp35:*
 137: em_estp36:*
 138: em_estp37:*
 139: em_estp38:*
 140: em_estp39:*
 141: em_estp40:*
 142: em_estp41:*
 143: em_estp42:*
 144: em_estp43:*
 145: em_estp44:*
 146: em_estp45:*
 147: em_estp46:*
 148: em_estp47:*
 149: em_estp48:*
 150: em_estp49:*
 151: em_estp50:*
 152: em_estp51:*
 153: em_estp52:*
 154: em_estp53:*
 155: em_estp54:*
 156: em_estp55:*
 157: em_estp56:*
 158: em_estp57:*
 159: em_estp58:*
 160: em_estp59:*
 161: em_estp60:*
 162: em_estp61:*
 163: em_estp62:*
 164: em_estp63:*
 165: em_estp64:*
 166: em_estp65:*
 167: em_estp66:*
 168: em_estp67:*
 169: em_estp68:*
 170: em_estp69:*
 171: em_estp70:*
 172: em_estp71:*
 173: em_estp72:*
 174: em_estp73:*
 175: em_estp74:*
 176: em_estp75:*
 177: em_estp76:*
 178: em_estp77:*
 179: em_estp78:*
 180: em_estp79:*
 181: em_estp80:*
 182: em_estp81:*
 183: em_estp82:*
 184: em_estp83:*
 185: em_estp84:*
 186: em_estp85:*
 187: em_estp86:*
 188: em_estp87:*
 189: em_estp88:*

190: qb_gss25:*
 191: qb_gss26:*
 192: qb_gss27:*
 193: qb_gss28:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	173.4	10.3	531	AA224131	AA224131:1 571 bp mRNA
2	157	9.3	782	AW134052	AW134052:1 111407 bp
3	155.4	9.2	641	AW76645	AW76645:1 388745 bp
4	115.2	6.8	611	AG719425	AG719425:1 5512 bp
5	99.8	5.9	908	BE546976	BE546976:1 601071950 bp
6	90.2	5.3	413	AW113677	AW113677:1 8019 bp
7	89.6	5.3	275	BA196485	BA196485:1 196485 bp
8	88.2	5.2	452	AA111286	AA111286:1 226403 bp
9	88.2	5.2	477	AA134113	AA134113:1 3384602 bp
10	88.2	5.2	479	AA153325	AA153325:1 24204 bp
11	88.2	5.2	549	AA153325	AA153325:1 24204 bp
12	86.4	5.1	395	AA153325	AA153325:1 24204 bp
13	86.4	5.1	399	AA153325	AA153325:1 24204 bp
14	86.4	5.1	399	AA153325	AA153325:1 24204 bp
15	86.4	5.1	446	AA153325	AA153325:1 24204 bp
16	86.4	5.1	465	AA153325	AA153325:1 24204 bp
17	85.2	5.1	450	AA153325	AA153325:1 24204 bp
18	85.2	5.1	645	AA153325	AA153325:1 24204 bp
19	84.8	5.0	384	AA153325	AA153325:1 24204 bp
20	84.8	5.0	384	AA153325	AA153325:1 24204 bp
21	84.8	5.0	390	AA153325	AA153325:1 24204 bp
22	84.8	5.0	428	AA153325	AA153325:1 24204 bp
23	84.8	5.0	451	AA153325	AA153325:1 24204 bp
24	84.8	5.0	509	AA153325	AA153325:1 24204 bp
25	84.8	5.0	574	AA153325	AA153325:1 24204 bp
26	84.8	5.0	588	AA153325	AA153325:1 24204 bp
27	84.8	5.0	655	AA153325	AA153325:1 24204 bp
28	84.2	5.0	156	AA153325	AA153325:1 24204 bp
29	84.2	5.0	192	AA153325	AA153325:1 24204 bp
30	84	5.0	955	AA153325	AA153325:1 24204 bp
31	83.6	4.9	714	AA153325	AA153325:1 24204 bp
32	82.8	4.9	594	AA153325	AA153325:1 24204 bp
33	82.2	4.9	317	AA153325	AA153325:1 24204 bp
34	82.2	4.9	431	AA153325	AA153325:1 24204 bp
35	82.2	4.9	1131	AA153325	AA153325:1 24204 bp
36	81.6	4.8	264	AA153325	AA153325:1 24204 bp
37	81	4.8	291	AA153325	AA153325:1 24204 bp
38	79.4	4.7	941	AA153325	AA153325:1 24204 bp
39	79.2	4.7	508	AA153325	AA153325:1 24204 bp
40	79.2	4.7	536	AA153325	AA153325:1 24204 bp
41	79.2	4.7	666	AA153325	AA153325:1 24204 bp
42	79.2	4.7	823	AA153325	AA153325:1 24204 bp
43	79	4.7	724	AA153325	AA153325:1 24204 bp
44	79	4.7	972	AA153325	AA153325:1 24204 bp
45	78.4	4.7	479	AA153325	AA153325:1 24204 bp

ALIGNMENTS

RESULT 1
 LOCUS AA224131/c
 DEFINITION 571 bp mRNA
 ACCESSION AA224131
 VERSION AA224131.1
 KEYWORDS EST
 SOURCE human

RESULT 15

LOCUS A1636264 446 bp mRNA EST 16-DEC-1999
 DEFINITION c29610.x1 NCI-CGAP_Kid1 Homo sapiens cDNA clone IMAGE:2296170 3' similar to qb:M1050 GUCCOXYOPTICOID RECEPTOR, BETA (HUMAN);, mRNA sequence.

ACCESSION A1636264 GI:4687594

VERSION EST

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Fukuyama; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS 1 (bases 1 to 446)

TITLE NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>

Journal National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

COMMENT Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher M. Skolnik, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.lbl.gov/bhrp/image/image.html

Insert length: 1206 Std Error: 0.00

Seq primer: 400P from Gibco.

Location/Qualifiers

1. 446

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2296170"

/clone_lib="NCI-CGAP_Kid1"

/db_host="db108"

/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with

a modified polylinker; Site_1: Not 1; Site_2: Eco RI;

Plasmid DNA from the normalized library NCI-CGAP_Kid3 was

prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(clones 132376-132391, 1456007-1456775, and

1500552-1502855). Subtraction by Bento Soares and M.

Falina Bonaldi."

BASE COUNT 113 a 97 c 105 g 131 t

ORIGIN

Query Match 5.1%; Score 86.4; DB 23; Length 446;

Best Local Similarity 64.5%; Pred. No. 2.5e-13;

Matches 129; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Db 220 TCTGGCTGCTGCTTCATGAGAGCTTCAGAGTTCATATGAGAGTTCATGAGAG 161

QY 596 GGTGAATGCTCTTTTAAAGAAAGATCATTGAGATGATGATGATGATGATGATGATGAT 655

Db 160 GCTGTAAGCTTTCTTCAAAAGACAGTGCAGACAGACAGACAGACAGTACCTATGCTGCGAG 101

QY 656 CCAATCAATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 715

Db 100 CCAATCAATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 41

QY 716 GATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785

Db 40 AATGCTTCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 21

Fri Apr 20 09:00:49 2001

us-08-906-365-1.rst

Page 11

456 TAAKALKEALDURENVTATNGTITDIFERKQVATLSEKCVFVAKVKGSPFERPG 225
 122 CEGGKATFKNSIGCHNYVCTATNGTITDIFERKQVATLSEKCVFVAKVKGSPFERPG 180
 226 YELVPGESADENGLPACAPKSGHAPRVRELLDALSPEDVLTLEAFPIVLTSP 285
 181 YRIVRGRSSSEQVHCLSKAKKNGHAPRVRELLDALSPEDVLTLEAFPIVLTSP 240
 286 SAPPTEASMMMSLTETKLADKEIVHMLGMAKIPGVFELSLDQVRLLESQMMVGLMW 345
 241 SMPTEASMMMSLTETKLADKEIVHMLGMAKIPGVFELSLDQVRLLESQMMVGLMW 300
 446 KSLDHPGKLIIFAPDLVLEDEKGVCELTLEFMDLATTSPPEKIQHKEYLVKAMLL 405
 301 KSLDHPGKLIIFAPDLVLEDEKGVCELTLEFMDLATTSPPEKIQHKEYLVKAMLL 360
 406 LNSSMYPLVATQVADSSSKRLAHLLNAVTDALVWVIAKSGISSQOOSMRLANILMLLSHV 465
 361 LNSSMYPLVATQVADSSSKRLAHLLNAVTDALVWVIAKSGISSQOOSMRLANILMLLSHV 420
 466 KHASNGKHEHLLMKCKRNVVYVYLLLEMLNAVHPRVGVCTGVSQVAPKPS*E6NSQ 525
 421 KHASNGKHEHLLMKCKRNVVYVYLLLEMLNAVHPRVGVCTGVSQVAPKPS*E6NSQ 480
 526 NPQSO 530
 1 111
 481 NLOSO 485

RESULT 2
 US-08-836-620A-2
 : Sequence 2, Application US/08836520A
 : Patent No. 5958710
 : GENERAL INFORMATION:
 : APPLICANT:
 : TITLE OF INVENTION: Orphan receptor
 : NUMBER OF SEQUENCES: 19
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: 05/08/836, 620A
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PC/EP96/03933
 : FILING DATE:
 : APPLICATION NUMBER: GB 951827.1
 : FILING DATE: 08-SEP-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 960550.4
 : FILING DATE: 15-MAR-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9607532.0
 : FILING DATE: 11-APR-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9609576.5
 : FILING DATE: 08-MAY-1996
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 485 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : ORIGINAL SOURCE:
 : ORGANISM: Rattus rattus
 : US-08-836-620A-2

Query Match 81.7%; Score 2291; DB 2; Length 485;
 Best Local Similarity 88.7%; Prid. No. 5,6e-237;
 Matches 410; Conservative 23; Mismatches 32; Indels 0; Gaps 0;

QY 46 MFEYSPAVMYVSTFSAVTHLEGGVCEGOTISPPVLTMPFCHSLVYVHPSLHVAEPKRS 105
 1 MFEYSPAVMYVSTFSAVTHLEGGVCEGOTISPPVLTMPFCHSLVYVHPSLHVAEPKRS 60
 QY 106 PWCFAASLEHTLPVNPETKRVKVSNGKASVPTGPKSKRIADPVAVNSVAGSYGOWS 165
 61 PWCFAASLEHTLPVNPETKRVKVSNGKASVPTGPKSKRIADPVAVNSVAGSYGOWS 120
 QY 166 CEGGKATFKNSIGCHNYVCTATNGTITDIFERKQVATLSEKCVFVAKVKGSPFERPG 225
 121 CEGGKATFKNSIGCHNYVCTATNGTITDIFERKQVATLSEKCVFVAKVKGSPFERPG 180
 QY 226 YELVPGESADENGLPACAPKSGHAPRVRELLDALSPEDVLTLEAFPIVLTSP 285
 181 YRIVRGRSSSEQVHCLSKAKKNGHAPRVRELLDALSPEDVLTLEAFPIVLTSP 240
 QY 286 SAPPTEASMMMSLTETKLADKEIVHMLGMAKIPGVFELSLDQVRLLESQMMVGLMW 345
 241 SMPTEASMMMSLTETKLADKEIVHMLGMAKIPGVFELSLDQVRLLESQMMVGLMW 300
 QY 446 KSLDHPGKLIIFAPDLVLEDEKGVCELTLEFMDLATTSPPEKIQHKEYLVKAMLL 405
 301 KSLDHPGKLIIFAPDLVLEDEKGVCELTLEFMDLATTSPPEKIQHKEYLVKAMLL 360
 QY 406 LNSSMYPLVATQVADSSSKRLAHLLNAVTDALVWVIAKSGISSQOOSMRLANILMLLSHV 465
 361 LNSSMYPLVATQVADSSSKRLAHLLNAVTDALVWVIAKSGISSQOOSMRLANILMLLSHV 420
 QY 466 KHASNGKHEHLLMKCKRNVVYVYLLLEMLNAVHPRVGVCTGVSQVAPKPS*E6NSQ 525
 421 KHASNGKHEHLLMKCKRNVVYVYLLLEMLNAVHPRVGVCTGVSQVAPKPS*E6NSQ 480
 526 NPQSO 530
 1 111
 481 NLOSO 485

RESULT 3
 US-08-836-620A-13
 : Sequence 13, Application US/08836520A
 : Patent No. 5958710
 : GENERAL INFORMATION:
 : APPLICANT:
 : TITLE OF INVENTION: Orphan receptor
 : NUMBER OF SEQUENCES: 19
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/836, 620A
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PC/EP96/03933
 : FILING DATE:
 : APPLICATION NUMBER: GB 9518272.1
 : FILING DATE: 08-SEP-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 960550.4
 : FILING DATE: 15-MAR-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9607532.0
 : FILING DATE: 11-APR-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9609576.5
 : FILING DATE: 08-MAY-1996
 : INFORMATION FOR SEQ ID NO: 13:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 484 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear

APPLICATION NUMBER: GB 9518272.1
 FILING DATE: 08-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9605550.4
 FILING DATE: 15-MAR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9607532.0
 FILING DATE: 11-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9609576.5
 FILING DATE: 08-MAY-1996
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 484 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: Mus musculus
 US-08-836-620A-14

Query Match 80.6%; Score 2262; DB 2; Length 484;
 Best Local Similarity 89.0%; Pred. No. 7,26-234;
 Matches 426; Conservative 23; Mismatches 35; Indels 0; Gaps 0;

QY 46 KILYSLAVNNTSLNSVMEGLLEGLSPNVLWPIPSHLSIVVHPLSHIYAFPSK 105
 DB 1 MAFYSAVMNYSVPSSTICQVWPTASPNVLTSPSLSLPLACOSLITAYPPSK 60
 QY 106 PWCENSLHTLVNPEETLKPKVSNPCASPVTPGSKPAPHCAPVSNVASYHGVMS 165
 DB 61 PWFASTSLHTLVNPEETLKPKLPSSTASPVSPSKLAPHCAPVSNVASYHGVMS 120
 QY 166 GGGGAAAFNNTSGHNVYCPALNQTLLGNPKRSYATLNRQYVYVWVYKQSKRLKQ 425
 DB 121 GGGGAAAFNNTSGHNVYCPALNQTLLGNPKRSYATLNRQYVYVWVYKQSKRLKQ 340
 QY 167 GGLVLELALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 285
 DB 161 YKLVLELALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 240
 QY 286 SALLLASMRSLALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 345
 DB 241 SALLLASMRSLALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 300
 QY 346 KSLDHRKLLFADLVLDHGRGCVGLDHLMLLATSKEEFTKIGHEVYVAMU 405
 DB 301 KSLDHRKLLFADLVLDHGRGCVGLDHLMLLATSKEEFTKIGHEVYVAMU 360
 QY 406 LNSMYELVATQDADSPKRLAHLLNAVTDALVWVIAKSGTSSQGSMPRLANLMLSHV 465
 DB 361 LNSMYELVATQDADSPKRLAHLLNAVTDALVWVIAKSGTSSQGSMPRLANLMLSHV 420
 QY 466 KASNSKMEHLNKRKYVYVLLLLDLNAVYKCKSKSTIGSESPEDSKREGSQ 525
 DB 421 KASNSKMEHLNKRKYVYVLLLLDLNAVYKCKSKSTIGSESPEDSKREGSQ 480
 QY 526 KEGS 529
 DB 481 NLOS 484

RESULT 6
 US-08-836-620A-15
 Sequence 15, Application US/08836620A
 Patent No. 5958710
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: Orphan receptor
 NUMBER OF SEQUENCES: 19
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPP)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/836,620A
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP96/01933
 FILING DATE:
 APPLICATION NUMBER: GB 9518272.1
 FILING DATE: 08-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9605550.4
 FILING DATE: 15-MAR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9607532.0
 FILING DATE: 11-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9609576.5
 FILING DATE: 08-MAY-1996
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 384 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-08-836-620A-15

Query Match 71.9%; Score 2017; DB 2; Length 384;
 Best Local Similarity 100.0%; Pred. No. 9,36-208;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LELVYHREGLLELALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 145
 DB 2 LELVYHREGLLELALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 61
 QY 146 AEECAVSLVASYHGVMSDEGKATFPESLGGMLYLTATRGCTLGNPKRSQACR 205
 DB 52 AEECAVSLVASYHGVMSDEGKATFPESLGGMLYLTATRGCTLGNPKRSQACR 121
 QY 206 LKRGYVAXVYGGSPFPYGLVYVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 265
 DB 122 LKRGYVAXVYGGSPFPYGLVYVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 181
 QY 266 LQVLELALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 325
 DB 182 LQVLELALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 241
 QY 326 DQVLELALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 385
 DB 242 DQVLELALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALALGEGALAL 301
 QY 386 APEELKIGHEVYVAMUINSMYELVATQDADSPKRLAHLLNAVTDALVWVIAKSG 445
 DB 302 APEELKIGHEVYVAMUINSMYELVATQDADSPKRLAHLLNAVTDALVWVIAKSG 361
 QY 446 ISSQGSMPRLANLMLSHVHNA 468
 DB 362 ISSQGSMPRLANLMLSHVHNA 384

RESULT 7
 US-08-836-620A-17
 Sequence 17, Application US/08836620A
 Patent No. 5958710
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: Orphan receptor
 NUMBER OF SEQUENCES: 19
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible


```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US-08-906-365-2
  FILING DATE:
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 101/EPO/03933
  FILING DATE:
  APPLICATION NUMBER: 08-9518272.1
  FILING DATE: 08-SEP-1995
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08-9518272.1
  FILING DATE: 15-MAR-1996
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08-96-07542.0
  FILING DATE: 11-APR-1996
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08-9609576.5
  FILING DATE: 08-MAY-1996
  INFORMATION FOR SEQ ID NO: 17:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 591 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
    ORGINAL SOURCE:
    ORGANISM: Homo sapiens
  US-08-906-620A-17

```

```

Query Match 44.0% Score 1222.5; Db 2; Length 591;
Best Local Similarity 48.1% Ident. No. 3,10-122;
Matches 262; Conserved 90; Mismatches 129; Indels 65; Gaps 14;

07 5 NSESSINSSSYN-SGSLLEH--GSLYESS--YVSHHFFAMTYSAVMVYSPS 60
08 1 NLEPLINP-----QKPLEKLEGYTLDSKPAVNYHGAVERNAANAAYG 74
09 61 NVNLEEDHCHQ-----ETSPNVLPFQHLSP-IVHROLSHLAE 101
10 75 Q-ILPEYHSGEAAFGSNLDGFFPLNSVSPSPMLLHPGULSPFQHPGGQPYLLE 133
11 102 TVRSHW-FAVSEFHTLVNPFETFEVNSNP-ASVYLDG-----SKGAHPVAVSD 154
12 134 NEEFNQTVHRAHAYVH--NSINPQDREKIASINIGSMAMPKAKELVCAVND 190
13 155 YASVHGVWSTGADKAPKSTGHNQYIFPATNCTIOPNPPSGVCLPKQVYCM 214
14 191 YASVHGVWSTGADKAPKSTGHNQYIFPATNCTIOPNPPSGVCLPKQVYCM 250
15 215 VQVSSPFFHGVVHVSQVSHV-THVATGAPRSTHAPVY-----IL 258
16 251 MFTLEHPRFSGMTHPRGCT-GLPEGVVSA-EMFAHLMWSTLMIFERKNS 305
17 259 LALALSTHGVVILFAPRPAVILSK--PSAETFEASGMMSLKLAKELVHM-SWAKK 316
18 306 LALSTHGVWSTGADKAPKSTGHNQYIFPATNCTIOPNPPSGVCLPKQVYCM 364
19 317 FGVELSTHGVVILFAPRPAVILSK--PSAETFEASGMMSLKLAKELVHM-SWAKK 376
20 365 FGVVILHGVVILFAPRPAVILSK--PSAETFEASGMMSLKLAKELVHM-SWAKK 424
21 427 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 435
22 435 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 494
23 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 495
24 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
25 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
26 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
27 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
28 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
29 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
30 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
31 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
32 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
33 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
34 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
35 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
36 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
37 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
38 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
39 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
40 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
41 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
42 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
43 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
44 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
45 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
46 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
47 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
48 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
49 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
50 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544

```

```

RESULT R
US-08-906-620A-16
Sequence 16, Alignment 95% (300/316)
Patent No. 5,958,710
GENERAL INFORMATION:
  APPLICANT:
  TITLE OF INVENTION: Update receptor
  NUMBER OF SEQUENCES: 1
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: 08-9609576.5
  FILING DATE:
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 101/EPO/03933
  FILING DATE:
  APPLICATION NUMBER: 08-9518272.1
  FILING DATE: 08-SEP-1995
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08-9609576.5
  FILING DATE: 08-MAY-1996
  INFORMATION FOR SEQ ID NO: 16:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 591 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
    ORGINAL SOURCE:
    ORGANISM: Rattus rattus
  US-08-906-620A-16

```

```

Query Match 44.0% Score 1214.5; Db 2; Length 591;
Best Local Similarity 48.1% Ident. No. 2,30-121;
Matches 262; Conserved 90; Mismatches 129; Indels 65; Gaps 13;

07 5 NSESSINSSSYN-SGSLLEH--GSLYESS--YVSHHFFAMTYSAVMVYSPS 60
08 1 NLEPLINP-----QKPLEKLEGYTLDSKPAVNYHGAVERNAANAAYG 74
09 61 NVNLEEDHCHQ-----ETSPNVLPFQHLSP-IVHROLSHLAE 101
10 75 Q-ILPEYHSGEAAFGSNLDGFFPLNSVSPSPMLLHPGULSPFQHPGGQPYLLE 133
11 102 TVRSHW-FAVSEFHTLVNPFETFEVNSNP-ASVYLDG-----SKGAHPVAVSD 154
12 134 NEEFNQTVHRAHAYVH--NSINPQDREKIASINIGSMAMPKAKELVCAVND 190
13 155 YASVHGVWSTGADKAPKSTGHNQYIFPATNCTIOPNPPSGVCLPKQVYCM 214
14 191 YASVHGVWSTGADKAPKSTGHNQYIFPATNCTIOPNPPSGVCLPKQVYCM 250
15 215 VQVSSPFFHGVVHVSQVSHV-THVATGAPRSTHAPVY-----IL 258
16 251 MFTLEHPRFSGMTHPRGCT-GLPEGVVSA-EMFAHLMWSTLMIFERKNS 305
17 259 LALALSTHGVVILFAPRPAVILSK--PSAETFEASGMMSLKLAKELVHM-SWAKK 316
18 306 LALSTHGVWSTGADKAPKSTGHNQYIFPATNCTIOPNPPSGVCLPKQVYCM 364
19 317 FGVELSTHGVVILFAPRPAVILSK--PSAETFEASGMMSLKLAKELVHM-SWAKK 376
20 365 FGVVILHGVVILFAPRPAVILSK--PSAETFEASGMMSLKLAKELVHM-SWAKK 424
21 427 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 435
22 435 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 494
23 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 495
24 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
25 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
26 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
27 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
28 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
29 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
30 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
31 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
32 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
33 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
34 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
35 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
36 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
37 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
38 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
39 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
40 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
41 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
42 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
43 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
44 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
45 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
46 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
47 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
48 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
49 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544
50 495 VSMV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV-ATVSHV 544

```



```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08-906-365-2
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02088
FILING DATE: 28-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93 110 298.2
FILING DATE: 28 JUN 1994
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Robert B.
REGISTRATION NUMBER: 22,980
REFERENCE/KEYWORD NUMBER: 6564-5019
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-5000
TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 773 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURES:
NAME/KEY: misc_location
LOCATION: (424-428)
OTHER INFORMATION: /note= "linker peptide."
FEATURE:
NAME/KEY: misc_location
LOCATION: (429-773)
OTHER INFORMATION: /note= "Estrogen binding domain."
US-08-564-264-1

```

```

Query Match: 28.9% Score 810; DB 3; Length 773;
Post local similarity: 30.9%; Pred. No. 8,56-78;
Matches: 108; Conservation: 41; Mismatches: 126; Gaps: 15;

5 NSTSTNSPSSYSGSSTLEFENSTLYPSGVDSHHEFYAFTEYSPAMNYSISNVTN 64
DB 278 NLVSTYNAALNNALVSTFAIKNS---FASHIGKH---LMTSP---LSMKGLTELIN 325
57 65 LRT-----GPKQKLTSPNWIWLPDRI SPIVVRQI SHLYAEQKSPWCARSLHTL 117
DB 426 VVGNSSSRKASAVARTYTHITAIPIHFALY-----SKYYAYD----- 365
57 118 FVNRETEKRVSGNRCASIV-----LQPSKRIAHCAVCS-- QYASQYHG 162
DB 506 PLSKMIAIKIPIIN---PIFFWGHLEQKSAKSPRIIPANNGITISQVLDTISY--- 418
57 164 VWSFHSVAFEPSTIGCHNATPATNCTTKSGPPSSQAPEIKRYTVGVWKGSSPPF 222
DB 419 -----INRSISVF-----GSMKGIKFD 436
57 223 KQTHLYVRSASDQDQATVAKPSQDAQVVP-----ELLILASIE 296
DB 437 KQTHLYVRSASDQDQATVAKPSQDAQVVP-----ELLILASIE 491
57 297 QIVVITLPAAPPAVITSP--PSAPTFASMMSTFTAIKEI VHMISWAKIIPGVFIS 324
DB 442 QNVVATLPAAPD IIVSYVPIPFPSFASMMSTFTAIKEI VHMISWAKIIPGVFIS 359
57 325 FQNVATITVWNEVIMMTMMSTFTFTNTIIPAPFVITFEGKQVEGILETFLMLAT 384
DB 551 HQVAVHILPAMIFIMINWVMSKRPVAILPAPNITIPNQGKRVESGVITFLMLATS 610
57 385 SPSRRIKIVRKRYIVKAMITINSMPRI VIAL--QVAGSSPSTAHINAVTIALWVIAF 443
DB 611 SPSRRIKIVRKRYIVKAMITINSMPRI VIAL--QVAGSSPSTAHINAVTIALWVIAF 678

```

```

57 444 SPSRRIKIVRKRYIVKAMITINSMPRI VIAL--QVAGSSPSTAHINAVTIALWVIAF 506
DB 671 ASHTEGVRQKAGITLISIRHRSKRQDHSVSKVQVAVVILHLLAHML 727

RESULT 12
US-08-693-940-3
Sequence 3, Application US/0869940
Patent No. 6136027
GENERAL INFORMATION:
APPLICANT: Yee, Jiling-Kuan
APPLICANT: Pichman, Theodore
APPLICANT: Chen, Shun-Tai
TITLE OF INVENTION: Inducible Expression System
TITLE OF INVENTION: Inducible Expression System
TITLE OF INVENTION: Inducible Expression System
TITLE OF INVENTION: Inducible Expression System
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: 285 Hamilton Ave., Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER REWARDABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,940
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Francis, Carol L.
REGISTRATION NUMBER: 46,513
TELEPHONE/TELEFAX NUMBER: 6510-055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDINESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-693-940-3

Query Match: 27.2% Score 742; DB 3; Length 651;
Post local similarity: 59.1%; Pred. No. 96-73;
Matches: 143; Conservation: 57; Mismatches: 48; Gaps: 11;

57 292 ALSPGATALLPAPPIVILSR--PFAV--LAVRMSSTLAKKAVHMSWAKIIPGVFIS 319
DB 313 LIL LIL LIL LIL LIL LIL LIL LIL LIL LIL LIL LIL LIL LIL LIL 311
57 315 SIIVAPVAVLDAVQD IIVSYVPIPFPSFASMMSTFTAIKEI VHMISWAKIIPGVFIS 423
DB 424 VVILHILVHLDAVILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHIL 483
57 423 VVILHILVHLDAVILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHIL 479
DB 424 VVILHILVHLDAVILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHIL 483
57 480 LILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHIL 438
DB 484 LILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHIL 643
57 483 VVILHILVHLDAVILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHIL 479
DB 484 VVILHILVHLDAVILHILHILHILHILHILHILHILHILHILHILHILHILHILHILHIL 483

```

```

Db      544  HMAKAGCILQOQHQLADLLILSHIRMSKGMHEHYSMKCKNVVPLDLEMLDAH 603
QY      499  VL 500
Db      604  RL 605

```

```

RESULT 14
US-08-836-620A-9
: Sequence 9, Application US/08/836,620A
: Patent No. 5958710
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Orphan receptor
: NUMBER OF SEQUENCES: 19
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/836,620A
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP96/03933
: FILING DATE:
: APPLICATION NUMBER: GB 9518272.1
: FILING DATE: 08-SEP-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9605550.4
: FILING DATE: 15-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9607532.0
: FILING DATE: 11-APR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9609576.5
: FILING DATE: 08-MAY-1996
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 243 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: ORIGINAL SOURCE:
: ORGANISM: Mus musculus
: US-08-836-620A-9

```

```

Query Match      23.8%; Score 668.5; DB: 2; Length 243;
Best Local Similarity 54.9%; Pred. No. 1.8e-63;
Matches 124; Conservative 52; Mismatches 49; Indels 1; Gaps 1;

QY      405  ELVHIMISAKKIPGVVELSLFDVQVHLSCWMEVIMKGLMMPISHPKTEFAPDLVLR 364
Db      1  ELVHIMINAKKVPGRGDLNHDQVHLLSCAMLELMIGLVMSMHKPKLLFAPRLDLDR 60
QY      465  DEKRVGVECTFLPMFLATTSRPFELKIQHKEYVLCVAMIIINSVYPLVATF QDADSS 423
Db      61  NQKRVPGMVEFLPMFLATSSRPFPMNIQGFPEVTKSTIIINSVYPLFESTLKSLPEK 120
QY      424  KRIAMLENAVIVAVIARAKSTISQAGSMFLANIMISIVHVASNKGMEILLNMKKN 483
Db      121  DHIRVILKIDILILHMAKAKNLTLDQOHRRLAQDLILSHIRMSNKGMEHLYNMCKN 180
QY      484  VVIVYDILLEMNAHVLPGKSSITIGSPSPAEISKSKESQNPQS 529
Db      181  VVPLVDLLEMLDMLRHAPASPMVPPPEPSQDGLATSSSTSAHS 226

```

```

RESULT 14
US-08-836-620A-8
: Sequence 8, Application US/08836620A
: Patent No. 5958710

```

```

: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Orphan receptor
: NUMBER OF SEQUENCES: 19
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/836,620A
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP96/03933
: FILING DATE:
: APPLICATION NUMBER: GB 9518272.1
: FILING DATE: 08-SEP-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9605550.4
: FILING DATE: 15-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9607532.0
: FILING DATE: 11-APR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9609576.5
: FILING DATE: 08-MAY-1996
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 243 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: ORIGINAL SOURCE:
: ORGANISM: Rattus rattus
: US-08-836-620A-8

```

```

Query Match      23.7%; Score 665; DB: 2; Length 243;
Best Local Similarity 56.9%; Pred. No. 3.4e-63;
Matches 124; Conservative 50; Mismatches 40; Indels 4; Gaps 2;

QY      305  ELVHIMISAKKIPGVVELSLFDVQVHLSCWMEVIMKGLMMPISHPKTEFAPDLVLR 364
Db      1  ELVHIMINAKKVPGRGDLNHDQVHLLSCAMLELMIGLVMSMHKPKLLFAPRLDLDR 60
QY      365  DEKRVGVECTFLPMFLATTSRPFELKIQHKEYVLCVAMIIINSVYPLVATF QDADSS 423
Db      61  NQKRVPGMVEFLPMFLATSSRPFPMNIQGFPEVTKSTIIINSVYPLFESTLKSLPEK 120
QY      424  KRIAMLENAVIVAVIARAKSTISQAGSMFLANIMISIVHVASNKGMEILLNMKKN 483
Db      121  DHIRVILKIDILILHMAKAKNLTLDQOHRRLAQDLILSHIRMSNKGMEHLYNMCKN 180
QY      484  VVIVYDILLEMNAHVLPGKSSITIGSPSPAEISKSKESQNPQS 529
Db      181  VVPLVDLLEMLDMLRHAPASPMVPPPEPSQDGLATSSSTSAHS 215

```

```

RESULT 15
US-08-836-620A-10
: Sequence 10, Application US/08836620A
: Patent No. 5958710
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Orphan receptor
: NUMBER OF SEQUENCES: 19
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/836,620A
: FILING DATE:

```

```

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: 96072796/9941
3 FILING DATE:
4 APPLICATION NUMBER: GB 961827.2.1
5 FILING DATE: 08 SEP 1995
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: GB 960550.4
8 FILING DATE: 15 MAR 1996
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: GB 9607532.0
11 FILING DATE: 11 APR 1996
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: GB 9609576.5
14 FILING DATE: 08 MAY 1996
15 INFORMATION FOR SEQ ID NO: 10:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 243 amino acids
18 TYPE: amino acid
19 TOPLOGY: linear
20 ORIGINAL SOURCE:
21 ORGANISM: Homo sapiens
22 DB: 08 836 620A-10

```

```

23077 Match: 23.54; Score 660.5; DB 2: Length 243;
23078 Best Local Similarity 60.48; Prod. No. 1.4e-62;
23079 Matches 119; Conservative 48; Mismatches 29; Indels 1; Gaps 1;

```

```

24 405 ELVIMISWARKTGVVLSLFDVGRLESCWVAVIMMLMWRSLDDGKLIFAHIVLDR 364
25 1 ELVIMIMWAKVPRVPLTLDGVHLEFAMIELMIGIWRSEHPRVRLFAVNLDR 60
26 465 DGRVWVLELEFEMIALISPEEHLKLNQVAVAMILLNSSMYPLVATGDADSS 423
27 61 NQKGVAVWVFPFMDATSSPRMMNCPPEVYKSTLNSDYLESLKSLER 120
28 424 EKLALHNAVVDALVVIKASGISDQGSRIANLMLSHVHASKKMEHLINRKN 483
29 121 DHHRVLFKTYTIDIMAVRNLITNNGVRIAGLILSLSHPRMSKQMEHLYSKCKN 180
30 484 VVIVVYDILLEMNAVL 500
31 181 VVHLYDILLEMNAHL 197

```

Search completed: April 19, 2001, 06:29:17
 Database: 25305, Size:

[Faint, illegible text covering the majority of the page, likely bleed-through from the reverse side.]

GeneCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

ORF protein: protein search using sw model

Run on: April 19, 2001, 09:24:56 : Search time 22.26 Seconds

(without alignments)
1616,582 Million cell updates/sec

111000
Percent Score: 2806
Sequence: 1 MLIKNSPSTLNSPSTNSVNSV.....EESFAESKSKESQNPUSQ 530

Search Index: HUSIM62
Gapop 10.0, Gapext 0.5

Search Pool: 195841 seqs, 6730655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum hit seq length: 0
Maximum hit seq length: 200000000

Post processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Hit description:
1: P1R1**
2: P1R2**
3: P1R3**
4: P1R4**

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is based on analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2905	100.0	530	2	J05939	estrogen receptor
2	2522	89.9	477	2	S71400	estrogen receptor
3	2284	81.4	503	2	JM0046	estrogen receptor
4	1260.5	43.9	589	1	U08106	estrogen receptor
5	1224.5	43.7	595	1	U08106	estrogen receptor
6	1219.5	43.5	595	2	J47140	estrogen receptor
7	1218.5	43.4	599	1	U08106	estrogen receptor
8	1216.5	43.4	600	1	U08106	estrogen receptor
9	1214	43.4	596	1	U08106	estrogen receptor
10	1199	42.7	620	2	J10423	estrogen receptor
11	1174.5	41.9	701	2	S64737	estrogen receptor
12	1144.5	40.4	535	2	AB0224	estrogen receptor
13	1120	39.9	524	2	A67177	estrogen receptor
14	641.5	22.5	433	2	S78382	estrogen receptor
15	619.5	22.1	433	2	S78382	estrogen receptor
16	597.5	21.4	521	2	A29445	estrogen receptor
17	471	16.6	467	2	A43761	estrogen receptor
18	469.5	16.7	923	2	A19596	estrogen receptor
19	465	16.6	923	2	A19596	estrogen receptor
20	459	16.4	446	2	A44181	estrogen receptor
21	459	16.4	446	2	A44181	estrogen receptor
22	458.5	16.3	533	2	A41651	estrogen receptor
23	457.5	16.3	758	2	S78382	estrogen receptor
24	456	16.3	448	2	S78382	estrogen receptor
25	456	16.3	448	2	S78382	estrogen receptor
26	455.5	16.2	520	2	A41778	estrogen receptor
27	453	16.1	848	2	S78382	estrogen receptor
28	453	16.1	940	2	A25923	estrogen receptor
29	452	16.1	463	2	S78382	estrogen receptor

40	452	15.1	463	2	B41727	retinoid X receptor
41	451	15.1	933	1	U08106	retinoid X receptor
42	447.5	15.0	501	1	U08106	retinoid X receptor
43	447.5	15.0	984	2	A29518	retinoid X receptor
44	446	15.9	441	2	S78382	retinoid X receptor
45	442.5	15.8	467	2	S78382	retinoid X receptor
46	442.5	15.8	467	2	A47278	retinoid X receptor
47	442.5	15.8	598	2	A47278	retinoid X receptor
48	442.5	15.8	786	2	A47278	retinoid X receptor
49	441.5	15.7	776	2	A47278	retinoid X receptor
40	441.5	15.7	981	2	A47278	retinoid X receptor
41	441	15.7	470	2	B41977	retinoid X receptor
42	441	15.7	563	2	B41977	retinoid X receptor
43	438.5	15.6	777	1	U08106	retinoid X receptor
44	438	15.6	462	1	S78382	retinoid X receptor
45	438	15.6	899	2	A47278	retinoid X receptor

ALIGNMENTS

RESULT 1

J05939

estrogen receptor beta - human

C:Species: Homo sapiens (man)

C:Date: 05 Feb 1999 #sequence_revision 05 Feb 1999 #seq_update 21 Jul 2000

C:Accession: J05939

R:Gadway, S.; Jinnou, S.; Watanabe, T.; Hiroki, H.; O'Noda, A.; Hosoi, T.; Oguchi, Y.; Mori

Biochem. Biophys. Res. Commun. 245: 122-126, 1998

A:Title: The complete primary structure of human estrogen receptor beta (hERbeta) and

A:Reference number: J05939; M010961 98078

A:Accession: J05939

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-530 (530A)

A:Cross reference: D001 AB006590, NID04201114; P10N:MAA249.1; P10N:Q2011142

A:Superfamily: estrogen receptor; cDNA transmembrane protein homology - ERB

P:147 408/500000: cDNA transmembrane protein homology - ERB

Query Match	100.0%	Score 2905	Id 47	Length 530
Best Local Similarity	100.0%	Prod. No. 1	Id 4711	
Matches	530	Conservative	02	Mismatches 02
QY	1	MLIKNSPSTLNSPSTNSVNSV	1	ESTR1
DB	1	MLIKNSPSTLNSPSTNSVNSV	1	ESTR1
QY	61	NTNLEKSPGRTISFNVLMET	61	ESTR1
DB	61	NTNLEKSPGRTISFNVLMET	61	ESTR1
QY	121	RTTLEKSPGRTISFNVLMET	121	ESTR1
DB	121	RTTLEKSPGRTISFNVLMET	121	ESTR1
QY	161	NTNLEKSPGRTISFNVLMET	161	ESTR1
DB	161	NTNLEKSPGRTISFNVLMET	161	ESTR1
QY	241	CAKAKKSPGRTISFNVLMET	241	ESTR1
DB	241	CAKAKKSPGRTISFNVLMET	241	ESTR1
QY	301	LADKELVHMSWAKKPGV	301	ESTR1
DB	301	LADKELVHMSWAKKPGV	301	ESTR1
QY	361	VLPPLCKVTEVTELEFME	361	ESTR1
DB	361	VLPPLCKVTEVTELEFME	361	ESTR1
QY	421	ISPKPAGTISNVVIA	421	ESTR1
DB	421	ISPKPAGTISNVVIA	421	ESTR1

Keywords: DNA binding; nucleus; phosphoprotein; steroid hormone receptor; transcribed

F1120/Estrogen: amino-terminal NH2

F121-209/Estrogen: DNA binding *status predicted <RNA>

F183-456/Estrogen: orba transforming protein homology - ERBA

F185-205/Estrogen: zinc finger CCGG motif

F121-245/Estrogen: zinc finger CCGG motif

F100-595/Estrogen: steroid binding *status predicted <RNA>

F120-305/Estrogen: site: phosphate (50%) (overlaid) *status predicted

Query Match

43.7% Score 1224.5; DB 1; Length 595;

Best Local Similarity 48.3%; Prot. No. 576-88;

Matches 263; Conservative 88; Mismatches 129; Indels 65; Gaps 14;

```

QY 5 NSPSSINSSYVNGSSGSLIPPH GSIVPSS--YVSHHETVAMTFYSFAVNNISPS 60
DB 21 NELEPLNRP-----QKTLERPLRDEYLDSEFAVNYVEAAELMAAANAAYVU 74
QY 61 NVLELNARDEG TTSNVAWPTGGASPEI VVHPTLSHYAE 101
DB 75 Q-TLLVYHNSFAAFPSNMLDHEPLNSVSPSLMLLHPRLQISPLQHGQGVITYLE 133
QY 122 LKQWV LAKVETETVHETLEFV LKQWV LKQWV LKQWV LKQWV LKQWV 154
DB 134 NHESSYVWFAVFAVYVPE--NENSPQDSELEAFSTNNGCSMAHSAKFTFYCAVND 193
QY 155 YASGHHYVWSSGEGFAFFESSTGNNYVGFALNPTLQFNKPSVAVGELKCYEVM 214
DB 191 YASGHHYVWSSGEGFAFFESSTGNNYVGFALNPTLQFNKPSVAVGELKCYEVM 250
QY 215 YVQSSPEPSSYVPLVVPKPSAVNPNKAKAPSSQANVAVP-----EL 258
DB 251 MKQSLKPKRQKAPKMKREPKLEF-----GESEFEGVACIMPAANMLPGLMTRSKNS 305
QY 259 LLDASPRQIVTTLFAPPRVILSR--PSAPTFEASMMSLTLAKRELVHMSAKKI 316
DB 306 LALSTLQWMSALIDAPRP--ILSEYDIPRPFSEASMMMLTLNADRELVHMMARKV 364
QY 417 PGEVEELSTEDVPRITFSPWMEVLMMGLMWSLIDHGLIFAPDLVDRDEKVEGLEI 376
DB 405 PGEVEELSTEDVPRITFSPWMEVLMMGLMWSLIDHGLIFAPDLVDRDEKVEGLEI 424
QY 427 PTEGLATLSPEPELKLQKRELVYKAMLLNNSMYPLVAT--QVADSSKRLAHLLNAVTD 435
DB 425 PTEGLATLSPEPELKLQKRELVYKAMLLNNSMYPLVAT--QVADSSKRLAHLLNAVTD 484
QY 436 ALVWVAKSLSSQSSKRLANLLMLSHVHASNKMEHLLNKKCNVAVVYLLLEML 495
DB 445 ALVWVAKSLSSQSSKRLANLLMLSHVHASNKMEHLLNKKCNVAVVYLLLEML 544
QY 496 NAHVL 508
DB 545 NAHVL 549

```

RESULT 6

147140

estradiol receptor pta

6 Spect: Sus scrofa domestica (domestic pig)

6 Spect: 21 Feb 1997 #sequence: revealed 21 Feb 1997 #rev: change 23 Aug 1999

6 Accession: 147140; S66250; S12402

6 Review imp: 0; Incomplete: F.W. 25-06, H.H.

6 Mol. Cell. Endocrinol. 104, 16-172, 1994

6 Title: The C-terminal half of the porcine estradiol receptor contains two post-translational

6 Reference number: 147140; M010:95080454

6 Accession: 147140

6 Status: preliminary; translated from OR/EMBL/DBD

6 Molecule type: mRNA

6 Accession: F-595 <RA>

6 Cross-reference: EMBL:247167; NID:9587554; FIDN:CA085524.1; PID:9587555

6 Experimental source: uterus

6 Method: H.H.; Moschler, L.; Jungblut, P.W.

6 Eur. J. Biochem. 231, 510-516, 1995

Active: Surface mapping of the ligand filled C-terminal half of the porcine estradiol

A Reference number: S66250; M010:95361877

A Accession: S66250

A Molecule type: Protein

A Accession: 297 307/310 313/329 325/329 342/347 349/317 420/405 473 THO

A Experimental source: uterus

6 Method: H.H. 120, 92 96, 1994

6 Title: Assignment of the ligand binding site of the porcine estradiol receptor to

6 Reference number: S12402; M010:94209484

6 Accession: S12402

6 Molecule type: protein

6 Accession: 307-323 <THO>

6 Support family: estrogen receptor; orba transforming protein homology

6 Keywords: DNA binding; nucleus; phosphoprotein; steroid hormone receptor; transcrip

F1120/Estrogen: amino-terminal NH2

F121-209/Estrogen: DNA binding *status predicted <RNA>

F183-456/Estrogen: orba transforming protein homology - ERBA

F185-205/Estrogen: zinc finger CCGG motif

F121-245/Estrogen: zinc finger CCGG motif

F156-271/Estrogen: nuclear location signal

F100-595/Estrogen: steroid binding *status predicted <RNA>

F121-227/237/240/Estrogen: site: phosphate (50%) (overlaid) *status predicted

F120-305/Estrogen: site: phosphate (50%) (overlaid) *status predicted

Query Match

43.5% Score 1219.5; DB 2; Length 595;

Best Local Similarity 46.0%; Prot. No. 140 87;

Matches 268; Conservative 89; Mismatches 126; Indels 99; Gaps 19;

```

QY 5 NSPSSINSSYVNGSSGSLIPPH GSIVPSS--YVSHHETVAMTFYSFAVNNISPS 56
DB 21 NELEPLNRP-----QKTLERPLRDEYLDSEFAVNYVEAAELMAAANAAYVU 60
QY 57 -----SLP-SNVNLEEGQPGHQ-----TISPVNLTPOHSLP- 88
DB 61 DFNAAAASAPVYGGSGIAYVGPSSFAAFGANGIQQPRLNSVSPVLLHPRQLSPV 120
QY 89 LVVHPTLSHYAEPLKSPK--FAKSLBITLPVNRITLKRKSKKCAKSPVGRG----- 141
DB 121 LHHQCGVIVYYLNNDSYAVFAVFAVYV--NSNBPQDSELEAFSTNNGCSMAHSAK 177
QY 142 SFAFAVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 201
DB 178 SAKETFYCAVNCNVASGHHYVWSSGEGFAFFESSTGNNYVGFALNPTLQFNKPS 247
QY 202 QAKREPKCYEVMVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 250
DB 248 QAKREPKCYEVMVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 292
QY 291 -----HAPVPELIDAI SPFGVITTLFAPPRVILSR--PSAPTFEASMMSLTKL 301
DB 293 PSPLLIKHKIKKNSPV--SLTADQMSIALIDAPRP--ILSEYDIPRPLSPASMMGLT 349
QY 302 ALKELVHMSAKKIITGVHSLFVHPLFSSQWMEVLMMGLMWSLIDHGLIFAPDLV 361
DB 350 ADRELVHMMNKAIRVGFPLSLHGVHLEFAMLETLMLDHWMSMEHPKRLDPAFNL 409
QY 362 LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV 469
DB 410 LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV LKQWV 495
QY 421 PTEGLATLSNAVVALVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 480
DB 470 PTEGLATLSNAVVALVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 529
QY 481 QKRVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 522
DB 530 QKRVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 567

```

RESULT 7

RESULT 9
 ORIGIN
 estrifrogen receptor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 28-Feb-1992 #Sequence:Levislon 14-Jul-1994 #text_change 15-Jun-1996
 C:Accession: A40907
 K:W:1997.1.0.1; Law, D.; Shapiro, D.J.
 M:1; Publication: 1, 355-362, 1987
 A:Title: The Xenopus laevis estrifrogen receptor - sequence homology with human and avian r

Query Match: 43.38; Score 121; DB 1; Length 586;
Best Local Similarity: 48.68; Field No. 3,Je-47;
Matches: 255; Conservative: 82; Mismatches: 138;
Indels: 50; Gaps: 12

RESULT 10
T10423
estrogen receptor - *Oryzias* sp. (strain d-TR)
C:Species: *Oryzias* sp.
A:Variety: strain d-TR
C:Date: 20-Sep-1999 #sequence=1
C:Accession: T10423
R:Okada, H.; Kawahara, T.; Yamashita, I.
submitted to the EMBL Nucleic Acids Library, March 1994
A:Description: Cloning of medaka estrogen receptor cDNA.
A:Reference number: Z17013

Query Match	42.7%	Score 1169	198.2	Length 620
Best Local Similarity	48.1%	Pident. No. 5,790-863		
Matches	265	Conservative	81	Mismatches 129
				Indels 46
				Gaps 16

07	Y	16	YVYSHUEY--FAWT-----EYUVAANYNYSP--ANTNLEGGEGKGTYSNVM 80
	
		62	FLEGTGYDAAPNATITPLYSOSTGYSAPILETNCHPSGSLQSGISGSR--ISPLAV 117
07	Y	81	PIPGHLSPIVNHQLSHYALYKSKWPLASLEHTLIVHETLRKVSGR--GASPV 137
	
		118	FSSPRLSPWM--HPSHHLETTSTUYKSS-----HOFASREDGSGREDTSGLE 167
07	Y	138	TGQPS-----KDAHFACSDVASCYHGYVSWSGEGRKAFKRSJGHNIMYICPANO 190
	
		168	LQAGADAGCEHEMAKTLRCAVCSIDYASGAYHVSCEGGRKAFKRSJGHNIMYICPANO 227
07	Y	191	CTTKRNPSSGACGLKCYEVCWKGSGRKGRCSTGYKYSGR-----SAD 236
	
		228	CTTFNHPHSLALCHLKHLYEVSNGZCYKRE--LEHLEH-- 265
07	Y	237	FQIHCAGAKKS--GSHAPVHEHLLDALSPQIVLTLAPPHVLLSP--SAPPEVA 292
	
		286	KVAFHDKRKRSTGAGAGGSGEPLSVTSIPREYVLLGCMH--LLSRGOKLRPEYEV 344
07	Y	293	SMKMSYLRADKELVHMLSMARKHGFELNLSFQVYVLLNSGMYVLMGIMHRSYHNG 352
	
		345	TMTTLTSMADFLVHMLAMAKLPLPGTQSLHDVLLSSMSFLVMIGLHNSHNG 404
07	Y	353	KLLEAFDLVLRDESGEGLTFPMLATSSGRTKLQGRKYTVFAMILLNSGMY 412
	
		405	KLIPQODILDRNEDGVYSGMTLEFDMILATASBEKVLKLRPEYVTLKALILLNSAPS 404
07	Y	413	LVTAYQD--ADSSKRLHLLNVTALVWVAKSSISQSGSMETANIMSHVPSKNS 471
	
		465	FCGTGMEPLHNSAAVOSMLDTITVALHYISQCYLAQDAPRAGVALLSHIRHNSK 524
07	Y	472	GMEHLHNKKNVNVYVLLLEMLNAVIV--NGKNS-----SITSGESVAF 516
	

A>Title: Full length sequence and in vitro expression of rainbow trout estrogen receptor
 A:Reference number: A27197; MBL:01006824
 A:Accession: A37197
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-574 <IAR>
 A:Cross-references: GB:X51454
 A:Keywords: DNA binding; nuclear; steroid hormone receptor; transcription regulation; zinc
 F145-427/Region: cDNA transforming protein homology - ERBA
 F147-167/Region: zinc finger
 F183-207/Region: zinc finger

Query Match 49.9%; Score 1120; DB 2; length 574;

Best Local Similarity 48.8%; Pred No. 1-80;

Matches 243; Conservative 70; Mismatches 127; Indels 58; Gaps 12;

```

QY 70 CQGTSPNVLPVTPGHSPLV---VHQQLSH-----LYAEPRKSWCARSLHNTLPVN 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 55 GSSPFGPFLVSSSPQLSPQLSPFLPPSHHGLPSQSYLETSLPPLYSVVNTOLSA 114
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 EELL---KEKVSNGCAVPTGPGSKELAEFCAGVSDVAGVHGVSCEGCKAFKRS 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 115 EKKICLASIKROOSVAGAGVAFEMANETRYCAVSDVAGVHGVSCEGCKAFKRS 174
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 177 ICHNNVITPATNCTTTFNPPKSGAVPLPGVEYDWKSGSPKPGCVPLVPR----- 231
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 179 ICHNNVITPATNCTTTFNPPKSGAVPLPGVEYDWKSGSPKPGCVPLVPR----- 233
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 232 -----GRANDFGHCAQKARSSG--HAPRVFELDALSPQVLTILE 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 234 GRAPPEPTVWSTQDQVQVPPSSSLNPGVWMPPT-----TPPPQ-VLFLQ 286
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 275 APPPIVITSR--SAPPTFASMMMSITKLADKELVHMTSMARKIDGVELSLDGQVILE 332
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 287 GQTP-AICSKQVARYETVIMTLLISMAKDELVHMTSMARKIDGVELSLDGQVILE 345
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 343 GTMMVETIMMIPSTPHGKTIIPAPVITDPHGGCVGCTITPMTLATSPPELKL 392
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 346 SSMLEVLMLGLWSTHCPKLTFAODLLDRSEVGVGMATPMTLATVSRFQMLKL 405
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 393 QREKELVAMILLNSSMYPLVT-ATQDASSKLAHLNAVDAVWIAKSGISSQOO 451
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 406 KPEEVCGLKAILLNGAFSPGNSVESLHNSAVESMLDITDALIHHSISGASVQO 465
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 463 SPSRLMLLMLSHVSHASRSPHILNRKTVNNAVITDLELNAHVLPCKRSITGSE 511
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 466 PRKQADLLLSHKIMSKKMBHYSIKCKNKPVYPLFLEMLDHPQ----- 515
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 512 CSTAHSKSGESGNTGS 529
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 516 -APKVAACGEGTGTGT 532
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14

B29345

steroid hormone receptor ERB2 precursor - human

C:Species: Homo sapiens (man)

C:Date: 15-Jan-1998 #sequence_revision 15 from 1988 #cgt_change 20-Sep-1999

C:Accession: B29345

C:Character: V. J. Yang, N. J. French, P. J. Evans, R. M.

Nature 311, 91-94, 1988

A>Title: Identification of a new class of steroid hormone receptors.

A:Reference number: A29345; MBL:83123546
 A:Accession: B29345
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-433 <GIS>
 A:Cross-references: GB:X51417; MBL:036610; FIDN:CA55779.1; FIDN:046611; GB:Y00290
 A:Keywords: unassigned extra-related proteins; cDNA transforming protein homology
 C:Keywords: DNA binding; steroid hormone receptor; transcription regulation; zinc fin
 F101 352/Region: cDNA transforming protein homology - ERBA
 F103 133/Region: zinc finger
 F133-163/Region: zinc finger

Query Match 52.5%; Score 615; DB 2; length 433;

Best Local Similarity 33.5%; Pred No. 9-42;

Matches 173; Conservative 78; Mismatches 147; Indels 119; Gaps 19;

```

QY 3 IKNSTSLNSPSYNSQSLELEHSLYPSYVPSHHPTAMISVAVMNSTPS 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EL 26 ITERS-----PSGIALSHHS-----PS 46
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 NVIN-PTGSGVLI-SQVWLPVPSQVVGSGVAVPQKSPW--CEAKSLHNTL 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 37 GSSDAGGFGMALGIHANTGSHMPAG-----AGAGNPKRSYHDIQSGMED-- 86
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 118 PNVPELKPYSVNP-----ASVW-PTGSGVAVPQKSPW--CEAKSLHNTL 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 87 -----SAKCEVNLNP-----KR---LALVQGLASVAVPQKSPW--CEAKSLHNTL 126
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 PPSSTQDQVQVPPSSSLNPGVWMPPT-----TPPPQ-VLFLQ 286
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 127 EKKICLASIKROOSVAGAGVAFEMANETRYCAVSDVAGVHGVSCEGCKAFKRS 186
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 231 KQSSNFGHCAQKARSSG--HAPRVFELDALSPQVLTILE 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 187 PRDSENSPY-----ISLQISPAKRELKLVSYLVAPPEKLYAMP 229
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 286 -SAPPTFASMMMSITKLADKELVHMTSMARKIDGVELSLDGQVILE 332
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 230 DVPPEGLKALITTCGLAPPELVPLISMAKDELVHMTSMARKIDGVELSLDGQVILE 289
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 345 WPSLHHPGKTIIPAPVITDPHGGCVGCTITPMTLATSPPELKL 392
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 290 YPSLPLDKLAVAEQIMDEHSLV-GLDLYPALQGVVPSKELVSRFQMLKL 405
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 405 IUNS-SMYPLVATGLAVSSEKLAHLNAVDAVWIAKSGISSQOO 451
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 349 LANSISM-----LENLEAVQKGLTFRALQVYPLFLEMLDHPQ----- 515
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 463 SHVRIASNGMEHLLNMCKKNVPRVYDLEMLNAHV 499
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 397 PLRQTAARAVOHFYSVKDQGVPMKLEPLEMAKV 433
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 15

estrogen receptor related protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1998 #sequence_revision 01 May 1996 #cgt_change 20-Sep-1996

C:Accession: S58087

R. Pettersson, K. Svensson, K. Mattsson, K. Carlsson, B. Carlsson, K. De-Kenstam,

submitted to the EMBL Data Library, July 1995

A:Description: Expression of a novel member of estrogen response element binding and

A:Reference number: S58087

A:Accession: S58087

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-433 <STP>

A:Cross-references: EMBL:X89594; MBL:0414855; FIDN:CA61755.1; FIDN:0414856

C:Keywords: unassigned extra-related proteins; cDNA transforming protein homology

C:Keywords: steroid hormone receptor; zinc finger

F101 352/Region: cDNA transforming protein homology - ERBA

Query Match 221% Score 619.50 DB 2: Length 433;

Best Local Similarity 44.0% Pval: No; 86-41;

Matches 170; Mismatches 66; Indels 144; Gaps 20;

```

07 4 1KNHGSINSSSTNSSTLLEHHSYITSSVDT--SHHEYPAMIFYSYAVMNYSTPS 60
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 10 1KTFSS-----SSSLALSHHS-----PS 36
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
09 61 NVINEDTNT-RQITSPNWMPTSHESPLVVRQLSHLYAEPOKSPW--CEARSLHTL 117
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 67 SSLSASSDEGIALSHAMILLSPREFAT-----AGIGGNCCKKSYEDCTSGIMED-- 86
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
07 118 EVNNELLKPKKYSNG-----ASIVTQCSREHAFPAVNSLYASGYHYCVMSGCCAKF 172
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 97 -SARKPEYMLNADP-----KR--LPLVCGDIASGYHYGVASCEACRAK 126
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
07 173 KPSKLEHNTYTFANNGTTPVNPFFSQATLKKCYEVSXVKQSSP--SEPGCYELVF 230
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 127 FPHLLGYNVNVYAINVETTF--P--GATPMPYVVMKIDSVPLKLVKQDQKPK 186
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
07 241 KQKSNAP-----GTPAGAFPSGSHAPVPELILIALSPECLVLLLEAEYHVLISPS 286
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 187 KRLLSENSPYINIPISPAKRP-----LKKVSNLIVEQ--DKLYAMPN-----D 231
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
07 287 APTFEASMMSLIKLAKELVHMLISMARRTPQFVELSLFDQVRLLESQNMVYLMGIMKR 346
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 242 LPSYIKKATITTCPLAQELVPLINMARHITPSPSTLIGDQMSLLQSAMMETLLICIVKR 291
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
07 447 SLDHWKCLFAPQVVLQDESKVNGILLEFLMLATTSREFEELKQHKVEYLCVAMILL 406
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 292 SLPYDRIKLAAYADYIMQEHSLVGLDILYKALIQIVRYKKLKVKEFEMLKALALA 350
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
07 407 NS-SMYPLVIALQADNSRKLALHLN-AVTIALVWVIKSGTSSQGSMTANLMLLSH 464
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 401 NSLSNMY-----TENLEAVQKLGILLHFAIDQVEL-----SQREHPRRACKLIMTLPL 398
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
07 465 VRIASNKIDREHLNNRCKNVVAVYDILLEMLNABV 499
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
10 409 LKQIDAKAVGHHVSNVSGVVMHPLLEMLLAAV 433
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

```

Search completed: April 19, 2001, 06:44:59
 Job Time: 144 sec

1

2

3

Matches: 489; Conservative: 229; Mismatches: 20; Gaps: 0; Gaps: 0;

27 1 MDIKRSGSSLSRSGTNGSSGSLIA LKNGSLIA LSSVDSHEDYVAMLPYSAAVNNSTPS 60
 1 MDIKRSGSSLSRSGTNGSSGSLIA LKNGSLIA LSSVDSHEDYVAMLPYSAAVNNSTPS 60
 61 NVNRLKGGTNGTSTGNTVAVPTTQHL SPVAVHQSLSLVAHKKSWCKAKSLPHLPVN 120
 61 NVNRLKGGTNGTSTGNTVAVPTTQHL SPVAVHQSLSLVAHKKSWCKAKSLPHLPVN 120
 61 SVNLEHGRPKQITSPMLNSTGHSLPLAVHQLSHLVAPKSWCEAPSEHLTPVS 120
 121 KETLKRVKSGNCAAGVGTGSKRAHEDVAVSDVASTVHVSWSJEPKAFKPSIQH 180
 121 KETLKRVKSGNCAAGVGTGSKRAHEDVAVSDVASTVHVSWSJEPKAFKPSIQH 180
 121 KETLKRVKSGNCAAGVGTGSKRAHEDVAVSDVASTVHVSWSJEPKAFKPSIQH 180
 181 NDVTCATNCTTCKNPKSGACPLPEYEVAVGVSPPPEPVPLVPGGNADQLH 240
 181 NDVTCATNCTTCKNPKSGACPLPEYEVAVGVSPPPEPVPLVPGGNADQLH 240
 241 CAGKAKSGCHAPRVRELLDALSPQIVLLEAPPHVLSRPSAPTEASMMSLTK 300
 241 CAGKAKSGCHAPRVRELLDALSPQIVLLEAPPHVLSRPSAPTEASMMSLTK 300
 401 LAKELVHMSMAKKIPGVELSLFDQVRLTESGMFVLMGIMWSTDPKLIAPDL 360
 401 LAKELVHMSMAKKIPGVELSLFDQVRLTESGMFVLMGIMWSTDPKLIAPDL 360
 361 VLEDEKKEVEGILEFEMLLATSPFEELKQREKLVKAMLLNSMAYLVATQGA 420
 361 VLEDEKKEVEGILEFEMLLATSPFEELKQREKLVKAMLLNSMAYLVATQGA 420
 421 DSKRKLHLLNAVTDALVWVIAKSGIISGQSGKRALMLLSHVHASKCKEHLINMK 480
 421 DSKRKLHLLNAVTDALVWVIAKSGIISGQSGKRALMLLSHVHASKCKEHLINMK 480
 481 CRNVAVVYDILLFMNAHVIFGKSSITSGSPADSKSGESQGNQDS 529
 481 CRNVAVVYDILLFMNAHVIFGKSSITSGSPADSKSGESQGNQDS 529

RESULT 2

09XSH5 PRELIMINARY; PRI: 527 AA.

AC 09XSH5; 01-NOV-1999 (TREMBLREL. 12, Created)
 DT 01-NOV-1999 (TREMBLREL. 12, last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, last annotation update)
 DE ESTROGEN RECEPTOR-BETA.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Sumnartia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRED HOLSTEIN; TISSUE-OVARIAN FOLLICLE;
 RX MEDLINE 99150196; PubMed 10626117.
 RA Kosenella C.S., Yuan X., Manikkam M., Calder M.D., Garverick H.A.,
 RA Lushin D.B.;
 RA "Cloning, sequencing, and localization of bovine estrogen receptor-
 beta within the ovarian follicle";
 RI -; SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: AF110402; AAD24432.1;
 DR HSSP: P03372; IHCP.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR001628;
 DR PFAM: PF00105; hormone_rec; 1.
 DR PRINTS: PR00047; STROIDEINCR.
 DR PRINTS: PR00450; VITAMINOR.
 DR PROSITE: PS00041; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;

Query Match: 87.8%; Score 2464.5; DR 6; Length 527;
 Best Local Similarity: 87.9%; Pred. No. 9, 96, 207;
 Matches: 466; Conservative: 22; Mismatches: 39; Indels: 3; Gaps: 1;

09 1 MDIKRSGSSLSRSGTNGSSGSLIA LKNGSLIA LSSVDSHEDYVAMLPYSAAVNNSTPS 60
 1 MDIKRSGSSLSRSGTNGSSGSLIA LKNGSLIA LSSVDSHEDYVAMLPYSAAVNNSTPS 60
 61 NVNRLKGGTNGTSTGNTVAVPTTQHL SPVAVHQSLSLVAHKKSWCKAKSLPHLPVN 120
 61 NVNRLKGGTNGTSTGNTVAVPTTQHL SPVAVHQSLSLVAHKKSWCKAKSLPHLPVN 120
 61 SVNLEHGRPKQITSPMLNSTGHSLPLAVHQLSHLVAPKSWCEAPSEHLTPVS 120
 121 KETLKRVKSGNCAAGVGTGSKRAHEDVAVSDVASTVHVSWSJEPKAFKPSIQH 180
 121 KETLKRVKSGNCAAGVGTGSKRAHEDVAVSDVASTVHVSWSJEPKAFKPSIQH 180
 121 KETLKRVKSGNCAAGVGTGSKRAHEDVAVSDVASTVHVSWSJEPKAFKPSIQH 180
 181 NDVTCATNCTTCKNPKSGACPLPEYEVAVGVSPPPEPVPLVPGGNADQLH 240
 181 NDVTCATNCTTCKNPKSGACPLPEYEVAVGVSPPPEPVPLVPGGNADQLH 240
 241 CAGKAKSGCHAPRVRELLDALSPQIVLLEAPPHVLSRPSAPTEASMMSLTK 300
 241 CAGKAKSGCHAPRVRELLDALSPQIVLLEAPPHVLSRPSAPTEASMMSLTK 300
 401 LAKELVHMSMAKKIPGVELSLFDQVRLTESGMFVLMGIMWSTDPKLIAPDL 360
 401 LAKELVHMSMAKKIPGVELSLFDQVRLTESGMFVLMGIMWSTDPKLIAPDL 360
 361 VLEDEKKEVEGILEFEMLLATSPFEELKQREKLVKAMLLNSMAYLVATQGA 420
 361 VLEDEKKEVEGILEFEMLLATSPFEELKQREKLVKAMLLNSMAYLVATQGA 420
 421 DSKRKLHLLNAVTDALVWVIAKSGIISGQSGKRALMLLSHVHASKCKEHLINMK 480
 421 DSKRKLHLLNAVTDALVWVIAKSGIISGQSGKRALMLLSHVHASKCKEHLINMK 480
 481 CRNVAVVYDILLFMNAHVIFGKSSITSGSPADSKSGESQGNQDS 529
 481 CRNVAVVYDILLFMNAHVIFGKSSITSGSPADSKSGESQGNQDS 529

RESULT 3

09T015 PRELIMINARY; PRI: 527 AA.

AC 09T015; 01-MAY-2000 (TREMBLREL. 13, Created)
 DT 01-MAY-2000 (TREMBLREL. 13, last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, last annotation update)
 DE ESTROGEN RECEPTOR-BETA.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Sumnartia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-OVARY;
 RC Carduus H., Higsby R.M., Pope W.F., Nephew K.L.;
 RI "Cloning and expression of estrogen receptor beta in the ovine
 ovary";
 RI Submitted (Aug-1999) to the EMBL/Genbank/DBJL databases.
 DR EMBL: AF177936; AAD5772.1;
 DR HSSP: P03372; IHCP.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR000536;
 DR INTERPRO: IPR001628;
 DR INTERPRO: IPR01723;
 DR PFAM: PF00105; hormone_rec; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;

01 MAY 2000 (ITERMUTED, 15, last sequence update)
 02 OCT 2000 (ITERMUTED, 15, last annotation update)
 03 ESTROGEN RECEPTOR BETA (EPASMENT)
 04 MACACA MULIATA (Rhesus macaque)
 05 EUKARYOTA: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 06 Mammalia: Eutheria: Primates: Catarrhini: Cercopithecoidea:
 07 Cercopithecoidea: Macaca.
 08 NCBI_TaxID: 9544.
 09 [1]
 10 SEQUENCE FROM N.A.
 11 WU X. X., MA X. H., SMITH G. S., NATHANIELS P. W.:
 12 "Polymorphic distribution of Epilypa and Epilypa MENA in intracellular
 13 tissues of the primate Rhesus monkey."
 14 Am. J. Physiol. 278:R199-R206(2000).
 15 EMBL: AF192249; AA054069.11.
 16 HSSV: F03472; IRE.
 17 INTERPRO: IPR000536; .
 18 INTERPRO: IPR001723; .
 19 INTERPRO: IPR001728; .
 20 PFAM: PF00104; Hormone_Foot_1.
 21 PRINTS: PR00046; STROBP-RE.
 22 PRINTS: PR00046; THRE-100-8MR.
 23 K00000.
 24 NCNLEK 1 1
 25 NCNLEK 279 279
 26 SEQUENCE 279 AA; 41105 MW; 858D9D7D01DAD301 CRC64;

Query Match 49.48% Score 1995.76 E-150 Length 279;

Best Local Similarity 96.28% E-02.144

Matches 274; Conserved 190; 4; Mismatches 2; Indels 0; Gaps 0;

01 279 GHHVAKKAPSSNAPVPELILALISPEQVLLLEAPPHVLLSRVAPTEASMMMS 297
 02 1 GHHVAKKAPSSNAPVPELILALISPEQVLLLEAPPHVLLSRVAPTEASMMMS 60
 03 279 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 457
 04 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 120
 05 458 PLVLLDDEKCVPTIFPMILATTSPPFTKIYKFFYGVKAMIIINSMPYLAT 417
 06 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 180
 07 418 GVALSSEKALINVTALVWVAKSGLSSGSMPLANILMLSHVHANKKMHLL 477
 08 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 240
 09 478 NNVVAVVILLEMLNANVLRSSSLQSSPAP 516
 10 478 NNVVAVVILLEMLNANVLRSSSLQSSPAP 279

RESULT 12
 279/082 PRELIMINARY: 181; 57% AA.

01 OCT 1999 (ITERMUTED, 12, created)
 02 OCT 1999 (ITERMUTED, 12, last sequence update)
 03 OCT 2000 (ITERMUTED, 15, last annotation update)
 04 ESTROGEN RECEPTOR TYPE BETA.
 05 SPERMATOPHYTES (Gymnosperms)
 06 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 07 Actinopterygii: Neopterygii: Teleostei: Euteleostei: Neuteleostei:
 08 Actinopterygii: Actinopterygii: Perciformes: Percoidae:
 09 Sparidae: Sparidae
 10 NCBI_TaxID: 9544.
 11 [1]
 12 SEQUENCE FROM N.A.
 13 TISSUE: Ovary
 14 SODERSTROM S., POWER D.M., OLSSON P.-E., CANARIO A.V.M.:
 15 "The estrogen receptors expressed in teleost fish: cDNA cloning,
 16 characterization and tissue distribution."

01 SUBMITTED (MAR 1999) to the EMBL/GenBank/NCBI databases
 02 SOURCE: ESTROGEN RECEPTOR NPTFAB (DB: ESTROGEN RECEPTOR)
 03 EMBL: AF146960; AA040341.11.
 04 HSSV: F03472; IRE.
 05 INTERPRO: IPR000536; .
 06 INTERPRO: IPR000536; .
 07 INTERPRO: IPR000536; .
 08 INTERPRO: IPR000536; .
 09 PFAM: PF00104; Hormone_Foot_1.
 10 PFAM: PF00104; 21-64; 1.
 11 PRINTS: PR00047; STROBP-RE.
 12 PRINTS: PR00047; STROBP-RE.
 13 PRINTS: PR00047; STROBP-RE.
 14 PRINTS: PR00047; STROBP-RE.
 15 NCNLEK 1 1
 16 NCNLEK 279 279
 17 SEQUENCE 279 AA; 42608 MW; 66612084B7D7D0 CRC64;

Query Match 49.48% Score 1995.76 E-150 Length 279;

Best Local Similarity 96.28% E-02.144

Matches 274; Conserved 190; 4; Mismatches 2; Indels 0; Gaps 0;

01 279 GHHVAKKAPSSNAPVPELILALISPEQVLLLEAPPHVLLSRVAPTEASMMMS 297
 02 1 GHHVAKKAPSSNAPVPELILALISPEQVLLLEAPPHVLLSRVAPTEASMMMS 60
 03 279 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 457
 04 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 120
 05 458 PLVLLDDEKCVPTIFPMILATTSPPFTKIYKFFYGVKAMIIINSMPYLAT 417
 06 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 180
 07 418 GVALSSEKALINVTALVWVAKSGLSSGSMPLANILMLSHVHANKKMHLL 477
 08 LLEAVVIMISWAKIPVVELSLFQVLESLWMEVLMGLMMSIHPPKILFA 240
 09 478 NNVVAVVILLEMLNANVLRSSSLQSSPAP 516
 10 478 NNVVAVVILLEMLNANVLRSSSLQSSPAP 279

RESULT 13
 279/082 PRELIMINARY: 181; 57% AA.

01 OCT 1999 (ITERMUTED, 12, created)
 02 OCT 1999 (ITERMUTED, 12, last sequence update)
 03 OCT 2000 (ITERMUTED, 15, last annotation update)
 04 ESTROGEN RECEPTOR TYPE BETA.
 05 SPERMATOPHYTES (Gymnosperms)
 06 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 07 Actinopterygii: Neopterygii: Teleostei: Euteleostei: Neuteleostei:
 08 Actinopterygii: Actinopterygii: Perciformes: Percoidae:
 09 Sparidae: Sparidae
 10 NCBI_TaxID: 9544.
 11 [1]
 12 SEQUENCE FROM N.A.
 13 TISSUE: Ovary
 14 SODERSTROM S., POWER D.M., OLSSON P.-E., CANARIO A.V.M.:
 15 "The estrogen receptors expressed in teleost fish: cDNA cloning,
 16 characterization and tissue distribution."

cDNA sequence of filoplia type beta estrogen receptor.
 BL Submitted (Oct-1996) to the FPMR/GenBank/DBU databases.
 -1- SUBMITTER LOCATION: NITELAR (BY SIMILARITY).
 LR FPMR: 076605; MAD00246.1; -
 LR HSS2: P04472; HRP.
 LR INFERRO: (P000424) -
 LR INFERRO: (P000566) -
 LR INFERRO: (P001628) -
 LR INFERRO: (P001724) -
 LR PRAM: (P00104) ; hormone_recep_1.
 LR PRAM: (P00105) ; 21-24; 1.
 LR PRINTS: P00047; SEROTININR.
 LR PRINTS: P00040; VITAMINR.
 LR PRINTS: P00098; STROPHOM_NER.
 LR PROSITE: P00041; NUCLEAR_RECEPTOR_1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger.
 SQ Sequence: 557 AA; 61779 MW; R2A74988C9G176 CPG64;

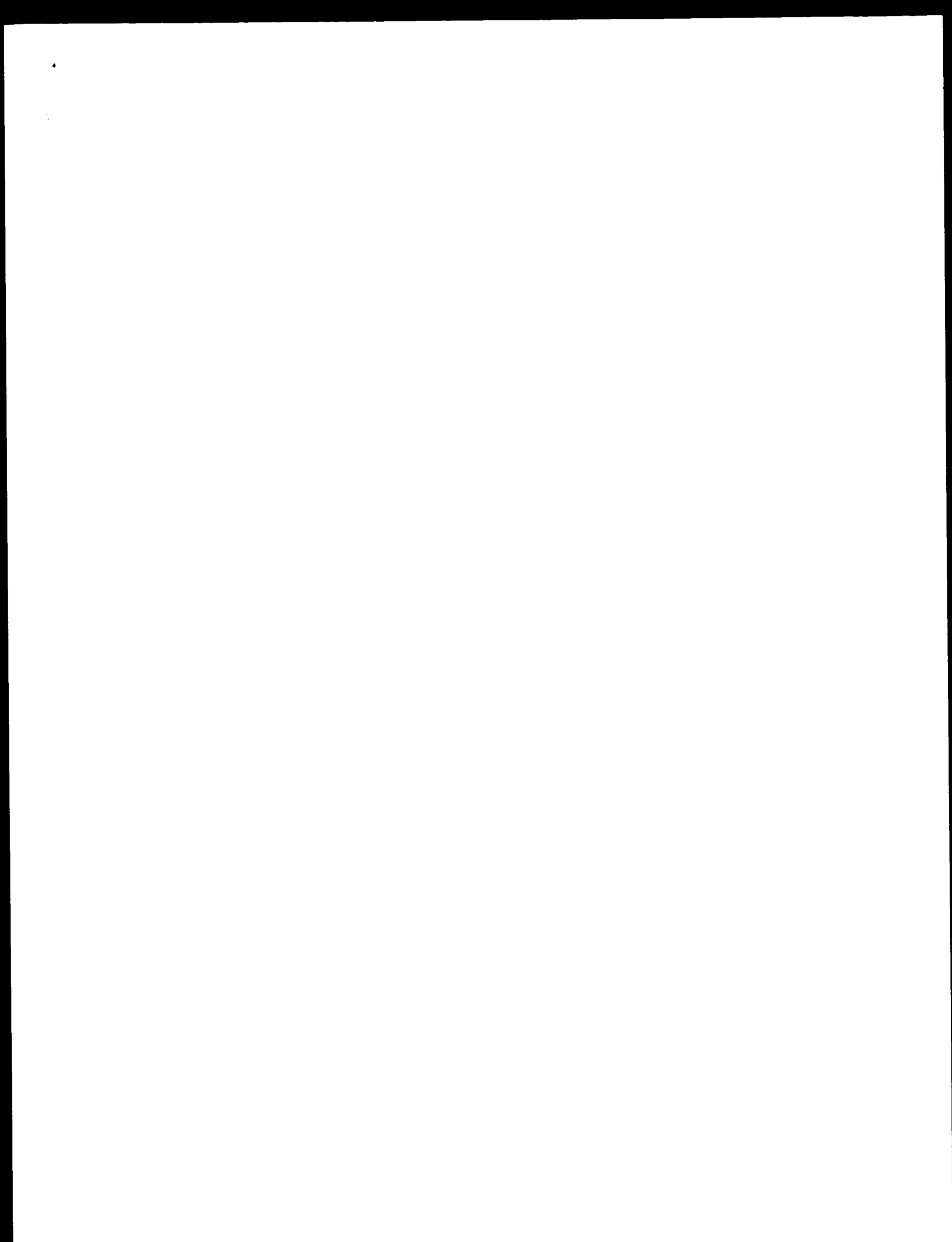
Query Match: 18.2%; Score: 135.2; DB: 13; Length: 557;

Best Local Similarity: 52.0%; Ident. No.: 11; E-109; Mismatches: 145; Gaps: 13;

Matches: 283; Conservative: 70; Models: 46; Gaps: 13;

1 MLEKNSPSLSNPSNYSNGLLEFHCSTYIPSSVDSHHEYPAMYPSPAMVNSTPS 60
 17 VSSSWANSPRLSPITSSPSLSHTSGPTLRSYTDLGHFTTIPYSSTISY---- 72
 61 NVNLEKADG - - - KQTLSPNVLWLPQGL SPVYH-PQLSHLYAEVPSHWCFAR 111
 73 GQPSLSPSSVQPSL SASI PWSQSHFVLPDLP HQGQPSQSQSQSQTFW---- 122
 112 STEHLVNRLELKKVSNRNASVYPRQSKRAHFCANSDVANSVHYGVWSCEGCKA 171
 123 -LSVLTLSKSVRR -SPESESWSASAKRAH dVAVGQASVHYGVWSCEGCKA 177
 172 FFKSLQGNVATQPAIMQCTLQRNKRKSVACRLKQCYWVMVQSPREPGYRLVPR 231
 179 FFKSLQGNVATQPAIMQCTLQRNKRKSVACRLKQCYWVMVQSPREPGYRLVPR 237
 242 QRSALQGLHAKAKPSGQHAQVREL--ELIALSPQGLVLTLEAPPHV-LISRPS 286
 248 KRLI KLSQSLQATGCTTPVAFSTNPPEFALPEQLEPILAEPELYLVDAK 295
 297 APPTFASMMSTLAKAGLVVMISWAKIAGVLSLFDVRLLESQMMEVLMQGLMR 346
 296 PELLFASVMMLLINAKRELVMISWAKIAGVLSLFDVRLLESQMMEVLMQGLMR 355
 347 SLQHSKELAFADLVDRERKVENGLLEIFDMIALISREKELNQRKRYGVAMQLL 406
 359 SVGLPQKLPQDLSREERNSVQGVHIFDMIALITRYRELKQREVEVGLKAMILL 415
 407 NSNR-YELVATLQFADSRKTAHILNAVITALVNVIAKSGLSNSQGSMLANILMLSHV 465
 416 NSNMQLSSQSPQGPSKSLKILDAMHAIYALGKQLTFPQDYLRLMLMLSHI 475
 466 PHANNEKHEHLINNRKENVAVYALILMLNAGVLCQKSLISQSPAE--LSKSKG 523
 476 KHYVNRKQNHQWQAMNINIVYVLLHMLAHIMH-----NSTLPHLPQGDQRKQ 527
 524 SQNP 527
 1 -
 106 528 SEVP 541

Search completed: April 19, 2001, 10:34:23
 Job time: 147.4 Sec



XX This sequence represents a novel oestrogen binding protein isolated from
 CC human testis cDNA in order to study upstream translation-initiation
 CC codons using 5' RACE-PCR technology. This receptor is able to bind and
 CC be activated by estradiol, estrone and estril, can be used in a screening
 CC assay for the identification of new drugs e.g. novel ligands of hormonal
 CC analogues.

XX Sequence 530 AA:

Query Match 100.0%; Score 241; DB 18; Length 530;
 Best Local Similarity 100.0%; Pred. No. 3e-23;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALKNSPSSINSNSPSSNSGSHLPLHGSLVITSSVDSHHEVPA 45
 DB 1 mdknspssinspsysncsgslplhgslvitypsvdsheypa 45

RESULT 2

W97858
 ID W97858 standard; Protein: 530 AA.

AC W97858;

DE 07-JUN-1999 (first entry)

DE Human oestrogen receptor-beta.

KW oestrogen receptor-beta; HER-beta; human.

OS Homo sapiens.

PN W09907847-A1.

PD 18-FEB-1999.

PF 20-JUL-1998; 98WO-US14944.

PR 05-AUG-1997; 97US-0906365.

FA (AMMP) AMERICAN HOME PROD CORP.

PI Bhat RA, Henderson RA, Hsiao C, Karathanasis SK;

DR WPI: 1999-167424/14.

N-PSDB: X24364.

PT Human oestrogen receptor-beta coding sequence - useful in the
 PT production of human oestrogen receptor-beta and identification of
 PT human oestrogen receptor-beta interactive compounds

PS Claim 17; Page 42-44; 49pp; English.

XX This polypeptide comprises a full-length human oestrogen receptor-beta
 CC (HER-beta), including 45 previously unknown N terminal amino acid
 CC residues that are believed to contribute to the transcription
 CC activation function of the receptor. The amino acid sequence was
 CC deduced from the nucleotide sequence (see X24364) of an isolated
 CC cDNA clone. HER-beta is selectively expressed in the thymus,
 CC spleen, ovary and testis. The invention encompasses HER-beta
 CC polypeptides and polypeptides, particularly peptides which
 CC include residues 1-45 of HER-beta. The invention also provides
 CC expression systems in which transcriptionally active HER-beta or
 CC fragments can be produced, as well as screening methods for
 CC identifying HER-beta agonists and antagonists (including
 CC tissue-specific oestrogens and anti-oestrogens), and HER-beta
 CC co-activators and inhibitors.

SO Sequence 530 AA:

Query Match 100.0%; Score 241; DB 20; Length 530;
 Best Local Similarity 100.0%; Pred. No. 4e-23;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALKNSPSSINSNSPSSNSGSHLPLHGSLVITSSVDSHHEVPA 45
 DB 1 mdknspssinspsysncsgslplhgslvitypsvdsheypa 45

RESULT 3

Y07270
 ID Y07270 standard; Protein: 548 AA.

AC Y07270;

DE 06-JUN-1999 (first entry)

DE Human oestrogen receptor.

OS Homo sapiens.

PN W09912961-A1.

PD 18-MAR-1999.

PF 04-SEP-1998; 98WO-US18577.

PR 29-MAY-1998; 98CB-0066932.

PR 08-SEP-1997; 97US-0058271.

PR 30-SEP-1997; 97US-0060520.

PR 30-OCT-1997; 97GB-0022884.

PA (MERL) MERCK & CO INC.

PI Wikinson H;

DR WPI: 1999-229232/19.

N-PSDB: X34309.

PT Estrogen receptor

PS Claim 1; Fig 1; 32pp; English.

XX This sequence represents a human oestrogen receptor. The receptor can
 CC be used to identify ligands that bind to human oestrogen receptor. The
 CC ligands can be used in a method for preventing or treating an oestrogen
 CC receptor mediated disease or condition, such as abnormal bone resorption,
 CC a cardiovascular disease, cancer, metabolic disorders, or central nervous
 CC system disorders. The ligand is especially used to treat osteoporosis,
 CC breast, uterine, ovarian or prostate cancer, diabetes or Alzheimer's
 CC disease.

SO Sequence 548 AA:

Query Match 100.0%; Score 241; DB 20; Length 548;
 Best Local Similarity 100.0%; Pred. No. 4e-23;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALKNSPSSINSNSPSSNSGSHLPLHGSLVITSSVDSHHEVPA 45
 DB 1 mdknspssinspsysncsgslplhgslvitypsvdsheypa 45

RESULT 4

W98127
 ID W98127 standard; Protein: 414 AA.

AC W98127;

XX Novel murine oestrogen receptor-beta genes and splice variants
 PI useful for treating oestrogen receptor-beta mediated disorders
 XX
 PS Claim 9; Fig 2b; 89pp; English.
 CC The present invention describes the murine complete oestrogen receptor
 CC beta (MER beta c); also designated mer beta-3). The present sequence
 CC represents MER beta-2, which is an isoform of MER beta c. The protein
 CC can be used to screen for agents that modulate or block the interaction
 CC between the ER beta c protein and physiological, stilbene or
 CC triphenylethylene (anti)oestrogens. The antibody which binds the protein
 CC can be used in assays to detect activation of the ER beta protein and
 CC can be used for subcellular localisation of the protein. The
 CC complementary nucleic acid can be used to inhibit expression of the
 CC ER beta c nucleic acid.
 XX
 SO Sequence 510 AA;
 Query Match 84.2%; Score 203; DB 20; Length 510;
 Best Local Similarity 84.4%; Pred. No. 2,4e-18;
 Matches 38; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 OY 1 MDKNSPSSLSNPSGSSSTPLEHSGTYPSSYVDSHHEPPA 45
 ID 20 meknspslsnpsgssstplehsgttypssyvdsheppa 64
 DB 20 meknspslsnpsgssstplehsgttypssyvdsheppa 64
 RESULT 7
 ID W98126 standard; Protein; 510 AA.
 AC W98126;
 DT 21-JUN-1999 (first entry)
 XX
 DE Mouse oestrogen receptor beta isoform MER beta-2.
 XX
 KW oestrogen receptor; ER beta-2; ER beta-c; mouse; tumour;
 KW diagnosis; therapy; isoform.
 XX
 OS Mus musculus.
 XX
 PN W09905170-A1.
 PD 04-FEB-1999.
 XX
 PE 28-JUL-1998; 98WO-0515549.
 XX
 PH 30-JUL-1997; 97US-0054210.
 PR 28-JUL-1997; 97US-0053869.
 XX
 PA (UYVA) UNIV YALE.
 XX
 PI Denton R, Fraslon C, Galien R, Ritsche-Rigon M;
 XX
 DE WPI: 1999 142852/12.
 XX
 PT Human oestrogen receptor beta-3 - useful in treating and monitoring
 PT tumours
 XX
 PS Example 2; Fig 2b; 89pp; English.
 CC The present sequence is alternatively spliced isoform mer beta 2
 CC of the murine complete oestrogen receptor-beta (ER beta-c, or
 CC MER beta-3). The MER beta-3 gene (see X24935) contains 9 exons,
 CC including the newly described exon 5B, which is not found in ER
 CC beta-1. Exon 5B is located in the ligand binding domain of MER
 CC beta-3, and probably plays a significant role in MER beta-3 function.
 CC The MER beta-2 clone was isolated from an embryonic stem cell
 CC genomic library. It lacks exon 3. 2 other isoforms, MER beta-1
 CC (see W98125) and rat ER beta-4 (see W98127), have also been

CC isolated, and the human exon 5B region (see X24934) has been
 CC identified. Based on these sequences, the invention provides
 CC methods for identifying agents that block or augment ER beta c
 CC mediated transcriptional regulation, methods to determine whether
 CC ER beta-c or its isoforms is being expressed in tissues or cells,
 CC and methods of identifying and using agents that block the
 CC transcriptional regulation of genes by ER beta c or its isoforms,
 CC which in turn modulates other biological and pathological
 CC processes. Gene therapy expression of ER beta c is envisaged.
 XX
 SO Sequence 510 AA;
 Query Match 84.2%; Score 203; DB 20; Length 510;
 Best Local Similarity 84.4%; Pred. No. 2,4e-18;
 Matches 38; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 OY 1 MDKNSPSSLSNPSGSSSTPLEHSGTYPSSYVDSHHEPPA 45
 ID 20 meknspslsnpsgssstplehsgttypssyvdsheppa 64
 DB 20 meknspslsnpsgssstplehsgttypssyvdsheppa 64
 RESULT 8
 ID Y04434 standard; Protein; 549 AA.
 AC Y04434;
 DT 25-JUN-1999 (first entry)
 XX
 DE Murine MER beta-1 clone protein sequence.
 XX
 KW Estrogen receptor beta; oestrogen receptor beta; ER beta.
 XX
 OS Mus sp.
 XX
 PN W09905171-A1.
 PD 04-FEB-1999.
 XX
 PE 28-JUL-1998; 98WO-0515540.
 XX
 PH 30-JUL-1997; 97US-0054210.
 PR 28-JUL-1997; 97US-0053869.
 XX
 PA (UYVA) UNIV YALE.
 XX
 PI Baron R, Chamblon P, Denton R, Dupont S, Garnier J;
 XX
 DE WPI: 1999-142853/12.
 XX
 PT Novel murine oestrogen receptor-beta genes and splice variants -
 PT useful for treating oestrogen receptor-beta mediated disorders
 XX
 PS Claim 9; Fig 2a; 89pp; English.
 CC The present invention describes the murine complete oestrogen receptor
 CC beta (MER beta c); also designated mer beta-3). The present sequence
 CC represents MER beta-1, which is an isoform of MER beta c. The protein
 CC can be used to screen for agents that modulate or block the interaction
 CC between the ER beta c protein and physiological, stilbene or
 CC triphenylethylene (anti)oestrogens. The antibody which binds the protein
 CC can be used in assays to detect activation of the ER beta protein and
 CC can be used for subcellular localisation of the protein. The
 CC complementary nucleic acid can be used to inhibit expression of the
 CC ER beta c nucleic acid.
 XX
 SO Sequence 549 AA;
 Query Match 84.2%; Score 203; DB 20; Length 549;
 Best Local Similarity 84.4%; Pred. No. 2,6e-18;
 Matches 38; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Fri Apr 20 09:00:56 2001

us-08-906-365-2_copy_1_45.rag

Database version 4.5
Copyright (c) 1994 - 2000 CompuGen Ltd.

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

Database: Protein search, using SW model

Query: us-08-906-365-2_copy_1_45
Search time 17.23 seconds
(without alignments)
46,899 Million cell updates/sec

ALIGNMENTS

Query Match	24.0%	Score	Seq. ID	Length	Pos
Best Local Similarity	40.0%	Prod. No. 572896	14	Matches	14
Conservative	5%	Mismatches	15%	Indels	1%

RESULT 1
US-08-468-036-5
Sequence 5, Appl 1

Patient No. 572896

GENERAL INFORMATION:

APPLICANT: DeMaggio, Anthony J.

APPLICANT: DeMaggio, Anthony J.

TITLE OF INVENTION: Materials and Methods for the Treatment of

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

US-08-468-036-5

DB 585 RLRNRSANSSQSLPQHTNNSVHNSFSYSHDE 619

RESULT 2

US-08-476-843-5

Sequence 5, Application US/08376843

Patent No. 5846764

GENERAL INFORMATION:

APPLICANT: De-Maggio, Anthony J.

APPLICANT: Hoskstra, Mel F.

TITLE OF INVENTION: Materials and Methods Relating to Proteins

TITLE OF INVENTION: that Interact with Casein Kinase I

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Looley, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 238 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/376,843

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/184,605

FILING DATE: 21-JAN-1994

ATTORNEY/AGENT INFORMATION:

NAME: No. 5846764and, Greta E.

REGISTRATION NUMBER: 35,302

REGISTERED-ATTORNEY: 23800/31784

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ. ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 668 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-476-843-5

Query Match 24.38; Score 58.5; DB 2; Length 668;

Best Local Similarity 40.0%; Prod. No. 7.5;

Matches 14; Conservation 5; Mismatches 15; Indels 1; Gaps 1;

QY 9 SANSNSYSNCSUSTIPF-HOSYIYHSSVNSHDE 42

DB 585 NLRNRSANSSQSLPQHTNNSVHNSFSYSHDE 619

RESULT 3

US-07-689-693B-4

Sequence 4, Application US/07689693B

Patent No. 5231011

GENERAL INFORMATION:

APPLICANT: David Hillyard

APPLICANT: Baldomero M. Olivera

TITLE OF INVENTION: Segregated Folding Determinants

TITLE OF INVENTION: for Small Disulfide-Rich Peptides

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Thorpe, No. 5231011th & Western

STREET: 9045 South 700 East, Suite 200

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage

COMPUTER: Compaq LTE/286

OPERATING SYSTEM: DOS 4.01

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/689,693B

FILING DATE: 19910418

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: none

FILING DATE: na

ATTORNEY/AGENT INFORMATION:

NAME: Westerd, M. Wayne

REGISTRATION NUMBER: 22,788

REFERENCE/DOCKET NUMBER: 9925

TELECOMMUNICATION INFORMATION:

TELEPHONE: (801) 566-6634

TELEFAX: (801) 566-0750

INFORMATION FOR SEQ. ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 27 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: CVTA Omega conotoxin.

IDENTIFICATION METHOD: Direct peptide

IDENTIFICATION METHOD: sequencing of purified Conus geographus

IDENTIFICATION METHOD: venom

US-07-689-693B-4

Query Match 42.8%; Score 55; DB 1; Length 27;

Best Local Similarity 57.98; Prod. No. 0.34;

Matches 11; Conservation 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 SPSSLSNPSNCSUSTIP 24

DB 3 SPSSCSPTSNSNCRSCNP 21

RESULT 4

US-07-689-693B-3

Sequence 3, Application US/07689693B

Patent No. 5231011

GENERAL INFORMATION:

APPLICANT: David Hillyard

APPLICANT: Baldomero M. Olivera

TITLE OF INVENTION: Segregated Folding Determinants

TITLE OF INVENTION: for Small Disulfide-Rich Peptides

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Thorpe, No. 5231011th & Western

STREET: 9045 South 700 East, Suite 200

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage

COMPUTER: Compaq LTE/286

OPERATING SYSTEM: DOS 4.01

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/689,693B

FILING DATE: 19910418

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: none

```

1 FILING DATE: na
2 ATTORNEY/AGENT INFORMATION:
3 NAME: WOSTER, M. Wayne
4 REGISTRATION NUMBER: 22,768
5 REFERENCE/KEY NUMBER: 9925
6 TELECOMMUNICATION INFORMATION:
7 TELEPHONE: (601) 566-0750
8 TELEFAX: (601) 566-0750
9 INFORMATION FOR SEQ ID NO: 1:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 74 amino acids
12 TYPE: AMINO ACID
13 TOPOLAGY: linear
14 MOLECULE TYPE: peptide
15 FEATURE:
16 NAME/KEY: prot protide sequence for four-loop
17 NAME/KEY: VIA: Gene Cloning from cDNA libraries.
18 IDENTIFICATION METHOD: Libraries were created
19 IDENTIFICATION METHOD: using oligo-dT primed pUC13 vector
20 US-08-624-123-4

```

```

Query Match 22.8% Score 552 DB 1: Length 74:
Host Local Similarity 57.9% Prod. No. 1,22:
Matches 11: Conservative 2: Mismatches 6: Indels 0: Gaps 0:

```

```

QY 6 SPSSINSTRYNSQSLIP 24
11 1 111111 11 1
DB 48 SPSSINSTRYNSQSLIP 66

```

```

RESULT 5
US-08-624-123-12
1 Sequence 12, Application US/9524124
2 Patent No. 579276
3 GENERAL INFORMATION:
4 APPLICANT: Shon, Ki-Joon
5 APPLICANT: Gilley, Michelle M.
6 APPLICANT: Gilley, Michelle M.
7 TITLE OF INVENTION: Coccolito peptides
8 NUMBER OF SEQUENCES: 14
9 CORRESPONDENT ADDRESS:
10 ADDRESSEE: Wobbe, Bart J., Howard & Civiletti
11 STREET: 1701 New York Avenue, N.W.
12 CITY: Washington
13 STATE: DC
14 COUNTRY: US
15 ZIP: 20005
16 COMPUTER RELEVABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 OPERATING SYSTEM: PC DOS/MS DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.40
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/9524124
22 FILING DATE: 07-04-1995
23 CLASSIFICATION: 546
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 08/424,561
26 FILING DATE: 17-APR-1995
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Shon, Jeffrey L.
29 REGISTRATION NUMBER: 26,957
30 REFERENCE/KEY NUMBER: 24260-107674-5
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 202-962-4810
33 TELEFAX: 202-962-8400
34 INFORMATION FOR SEQ ID NO: 12:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 74 amino acids

```

```

1 TYPE: amino acid
2 STRANDEDNESS: single
3 TOPOLOGY: linear
4 MOLECULE TYPE: protein
5 HYDROPHILIC: No
6 US-08-624-123-12

```

```

Query Match 22.8% Score 552 DB 1: Length 74:
Host Local Similarity 57.9% Prod. No. 1,22:
Matches 11: Conservative 2: Mismatches 6: Indels 0: Gaps 0:

```

```

QY 6 SPSSINSTRYNSQSLIP 24
11 1 111111 11 1
DB 48 SPSSINSTRYNSQSLIP 66

```

```

RESULT 6
PCT-US96-05262-14
1 Sequence 14, Application PCT/US96-05262
2 GENERAL INFORMATION:
3 APPLICANT: Shon, Ki-Joon
4 APPLICANT: Gilley, Michelle M.
5 APPLICANT: Gilley, Michelle M.
6 APPLICANT: Yoshikawa, Junji
7 APPLICANT: Marsh, Karen
8 APPLICANT: Cruz, Leonard J.
9 APPLICANT: H. Hara, David R.
10 TITLE OF INVENTION: Coccolito peptides
11 NUMBER OF SEQUENCES: 14
12 CORRESPONDENT ADDRESS:
13 ADDRESSEE: Wobbe, Bart J., Howard & Civiletti, LLP
14 STREET: 1701 New York Avenue, N.W., Suite 7000
15 CITY: Washington
16 STATE: DC
17 COUNTRY: U.S.A.
18 ZIP: 20005
19 COMPUTER RELEVABLE FORM:
20 MEDIUM TYPE: Floppy disk
21 OPERATING SYSTEM: PC DOS/MS DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.40
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: PCT/US96-05262-14
25 FILING DATE: 17-APR-1996
26 CLASSIFICATION:
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: US 08/424,561
29 FILING DATE: 17-APR-1995
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Shon, Jeffrey L.
32 REGISTRATION NUMBER: 26,957
33 REFERENCE/KEY NUMBER: 24260-107674
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: 202-962-4848
36 TELEFAX: 202-962-8400
37 INFORMATION FOR SEQ ID NO: 14:
38 SEQUENCE CHARACTERISTICS:
39 LENGTH: 74 amino acids
40 TYPE: amino acid
41 STRANDEDNESS: single
42 TOPOLOGY: linear
43 MOLECULE TYPE: protein
44 HYDROPHILIC: No
45 PCT-US96-05262-14

```

```

Query Match 22.8% Score 552 DB 1: Length 74:
Host Local Similarity 57.9% Prod. No. 1,22:
Matches 11: Conservative 2: Mismatches 6: Indels 0: Gaps 0:
QY 6 SPSSINSTRYNSQSLIP 24
11 1 111111 11 1

```

Db 48 SPANSCPTSYNCRSNCNP 66

RESULT 7

US-08-195-152-2

Sequence 2, Application US/08195152

Patent No. 5673,541

GENERAL INFORMATION:

APPLICANT: Bontol, Nancy M.

APPLICANT: Leiserson, William M.

TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST

TITLE OF INVENTION: PROTEINS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/195,152

FILING DATE: 14-FEB-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Trecartin, Richard E.

REGISTRATION NUMBER: 31,801

REFERENCE/PROCKET NUMBER: A-57551,351,3MS

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 760 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-195-152-2

Query Match 22.8%; Score 55; DB 1; Length 760;

Best Local Similarity 40.2%; Pred. No. 26;

Matches 13; Conservative 8; Mismatches 14; Indels 8; Gaps 1;

QY 9 SLKSTSTNGSGSLDLEHGSIVYHPSYV-----DSHHEV 43

Db 220 SVAGSNLYGSSASNPDLKGLAVVSSAVAAAAAAYDDKHVD 262

RESULT 8

US-08-472-934-4

Sequence 4, Application US/08472934

Patent No. 5753446

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL

TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lahlive and Cockfield

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,934

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/440,421

FILING DATE: 15-May-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/354,516

FILING DATE: 21-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 5,405,941

FILING DATE: 15-Apr-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/324,460

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/11690

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/04178

FILING DATE: 15-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Giulio A. DeConti, Jr. Esq.

REGISTRATION NUMBER: 31,503

REFERENCE/PROCKET NUMBER: CFI-0942V-72

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 619 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-472-934-4

Query Match 22.0%; Score 53; DB 1; Length 619;

Best Local Similarity 36.5%; Pred. No. 47;

Matches 19; Conservative 5; Mismatches 18; Indels 10; Gaps 4;

QY 1 MDIKNSPSSLSNP---SSYNGSGSLDLEHGSIVYHPS-----SYVDSHHEV 43

Db 197 MDMLDPLSLSSPNSGSGSGSLDPLD-gpeyprpmuPAQSYVHHDGDF 242

RESULT 9

US-08-472-934-12

Sequence 12, Application US/08472934

Patent No. 5753446

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL

TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lahlive and Cockfield

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

```

1 APPLICATION NUMBER: US-08/423,934
2 FILING DATE: 06-JUN-1995
3 CLASSIFICATION: 435
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: US-08/440,421
6 FILING DATE: 15-MAY-1995
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: 08/454,516
9 FILING DATE: 21-FEB-1995
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: US-9,405,941
12 FILING DATE: 15-APR-1994
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: 08/423,460
15 FILING DATE: 14-OCT-1994
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: PCT/US94/11690
18 FILING DATE: 14-OCT-1994
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: PCT/US94/04178
21 FILING DATE: 15-APR-1994
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Giulio A. DeGottis, Jr., Esq.
24 REGISTRATION NUMBER: 41,504
25 REFERENCE/WORK KEY NUMBER: CPT-004DVC12
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (617) 227-7400
28 TELEFAX: (617) 227-9941
29 INFORMATION FOR SEQ ID NO: 1:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 619 amino acids
32 TYPE: amino acid
33 Topology: linear
34 MULTIPLE TYPE: protein
35 US-08-472,934-12
36
37 Query Match: 42.0%, Score 54; DB 1; Length 619;
38 Host Local Similarity: 46.5%; Prod. No. 47;
39 Matches: 19; Conservative: 5; Mismatches: 18; Indels: 10; Gaps: 3;
40
41 RESULT 1
42 US-08-423,934-4
43 Sequence 4, Application US/0842460A
44 Patent No. 5054043
45 GENERAL INFORMATION:
46 APPLICANT: JOHNSON, GARY L.
47 TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CHL.
48 TITLE OF INVENTION: RESISTIVENESS TO EXTERNAL SIGNALS
49 NUMBER OF SEQUENCES: 10
50 CORRESPONDENCE ADDRESS:
51 ADDRESSEE: SHERIDAN ROSS & MCINTOSH
52 STREET: 1700 LINCOLN STREET, SUITE 4500
53 CITY: DENVER
54 STATE: CO
55 COUNTRY: USA
56 ZIP: 80204
57 COMPUTER READABLE FORM:
58 MEDIUM TYPE: floppy disk
59 COMPUTER: IBM PC compatible
60 OPERATING SYSTEM: PC-DOS/MS-DOS
61 SOFTWARE: Patcut to Release #1.0, Version #1.25
62 CURRENT APPLICATION DATA:
63 APPLICATION NUMBER: 08/423,934
64 FILING DATE: 14-OCT-1994
65 PRIOR APPLICATION DATA:
66 APPLICATION NUMBER: 08/423,460
67 FILING DATE: 14-OCT-1994
68 PRIOR APPLICATION DATA:
69 APPLICATION NUMBER: PCT/US94/04178
70 FILING DATE: 15-APR-1994
71 ATTORNEY/AGENT INFORMATION:
72 NAME: KARA, Gail Ho J.
73 REGISTRATION NUMBER: 44,106
74 REFERENCE/WORK KEY NUMBER: CPT-004DVC12
75 TELECOMMUNICATION INFORMATION:
76 TELEPHONE: (617) 227-7400
77 TELEFAX: (617) 227-9941
78 INFORMATION FOR SEQ ID NO: 4:

```

```

1 FILING DATE: 14-Apr-1994
2 ATTORNEY/AGENT INFORMATION:
3 NAME: KOWALIK, JOSEPH E.
4 REGISTRATION NUMBER: 43,005
5 REFERENCE/WORK KEY NUMBER: 2879-1-1
6 TELECOMMUNICATION INFORMATION:
7 TELEPHONE: 603/663,9700
8 TELEFAX: 603/663-0224
9 INFORMATION FOR SEQ ID NO: 4:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 619 amino acids
12 TYPE: amino acid
13 Topology: linear
14 MULTIPLE TYPE: protein
15 US-08-423,934-4
16
17 Query Match: 42.0%, Score 54; DB 1; Length 619;
18 Host Local Similarity: 46.5%; Prod. No. 47;
19 Matches: 19; Conservative: 5; Mismatches: 18; Indels: 10; Gaps: 3;
20
21 RESULT 1
22 US-08-461,146-4
23 Sequence 4, Application US/08461146
24 Patent No. 5081205
25 GENERAL INFORMATION:
26 APPLICANT: JOHNSON, GARY L.
27 TITLE OF INVENTION: METHOD FOR REGULATING MICROBIAL ACTIVITY
28 NUMBER OF SEQUENCES: 14
29 CORRESPONDENCE ADDRESS:
30 ADDRESSEE: Lattive and Cockfield
31 STREET: 69 State Street
32 CITY: Boston
33 STATE: Massachusetts
34 COUNTRY: U.S.A.
35 ZIP: 02109
36 COMPUTER READABLE FORM:
37 MEDIUM TYPE: floppy disk
38 COMPUTER: IBM PC compatible
39 OPERATING SYSTEM: PC-DOS/MS-DOS
40 SOFTWARE: Patcut to Release #1.0, Version #1.25
41 CURRENT APPLICATION DATA:
42 APPLICATION NUMBER: 08/461,146
43 FILING DATE: 05-JUN-1995
44 CLASSIFICATION: 435
45 PRIOR APPLICATION DATA:
46 APPLICATION NUMBER: US-08/454,516
47 FILING DATE: 21-FEB-1995
48 PRIOR APPLICATION DATA:
49 APPLICATION NUMBER: US-08/494,274
50 FILING DATE: 15-APR-1994
51 PRIOR APPLICATION DATA:
52 APPLICATION NUMBER: 08/423,460
53 FILING DATE: 14-OCT-1994
54 PRIOR APPLICATION DATA:
55 APPLICATION NUMBER: PCT/US94/04178
56 FILING DATE: 15-APR-1994
57 ATTORNEY/AGENT INFORMATION:
58 NAME: KARA, Gail Ho J.
59 REGISTRATION NUMBER: 44,106
60 REFERENCE/WORK KEY NUMBER: CPT-004DVC12
61 TELECOMMUNICATION INFORMATION:
62 TELEPHONE: (617) 227-7400
63 TELEFAX: (617) 227-9941
64 INFORMATION FOR SEQ ID NO: 4:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 619 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-146C-4

Query Match 22.0%; Score 53; DB 2, Length 619;
Best Local Similarity 36.5%; Pred. No. 37;
Matches 19; Conservative 5; Mismatches 18; Indels 10; Gaps 3;
QY 1 MDIKNSPSSLSNP---SSYNGSOSIPLLEHGSITYPS-----SYVDSHHEV 43
DB 197 MDQMIDPLSLSSPENSNGSGSPSLSPID-GEVYPKSMPAPAGVPPNDHGEF 247

RESULT 12
US-08-461-146C-12
Sequence 12, Application US/08461146C
Patent No. 5981265
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Lohive and Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,146C
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/354,516
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,254
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: KARA, Catherine J.
REGISTRATION NUMBER: P41,106
REFERENCE/DOCKET NUMBER: CPT-004CNI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 619 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-146C-12

Query Match 22.0%; Score 53; DB 2, Length 619;

Best Local Similarity 36.5%; Pred. No. 37;
Matches 19; Conservative 5; Mismatches 18; Indels 10; Gaps 3;
QY 1 MDIKNSPSSLSNP---SSYNGSOSIPLLEHGSITYPS-----SYVDSHHEV 43
DB 197 MDQMIDPLSLSSPENSNGSGSPSLSPID-GEVYPKSMPAPAGVPPNDHGEF 247

RESULT 13
US-08-461-145C-4
Sequence 4, Application US/08461145C
Patent No. 6074861
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: NOVEL MEKK PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Lohive and Cockfield, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,145C
FILING DATE: 5-JUNE-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,254
FILING DATE: 11-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/354,516
FILING DATE: 21-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: P-41,106
REFERENCE/DOCKET NUMBER: CPT-004CNI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 619 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-145C-4

Query Match 22.0%; Score 53; DB 4; Length 619;
Best Local Similarity 36.5%; Pred. No. 37;
Matches 19; Conservative 5; Mismatches 18; Indels 10; Gaps 3;

QY 1 MDIKNSPSSLSNP---SSYNGSOSIPLLEHGSITYPS-----SYVDSHHEV 43
DB 197 MDQMIDPLSLSSPENSNGSGSPSLSPID-GEVYPKSMPAPAGVPPNDHGEF 247

RESULT 14

Fri Apr 20 09:00:56 2001

us-08-906-365-2_copy_1_45.ra1

RESULT 14
S73697
probable X-pro dipeptidase (EC 3.4.13.9) pepX_Mycoplasma pneumoniae (strain ATCC 29342)
N:Alternate names: hypothetical protein P01_C01354
C:Species: Mycoplasma pneumoniae
A:Varianty: ATCC 29342
E:Updated: 10 Sep 1999 #sequence_level: 10 Sep 1999 #ext change: 21-Jul-2000
C:Accession: S73697
P:Immolation: K. Hultbert, H. Placens, H. Pirkel, E. Li, B.C. Herrmann, R. Nucleic Acids Res 23, 4420-4449, 1996
M:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A:Contents: annotation, determination, confirmed by chemical synthesis
A:Note: disulfide bonds determined and confirmed by chemical synthesis
R:Davis, J.H.; Bradley, F.K.; Miljanich, G.P.; Nadashi, I.; Ramchandran, J.; Basus, R.Davis, J.H.; Bradley, F.K.; Miljanich, G.P.; Nadashi, I.; Ramchandran, J.; Basus, submitted to the Brookhaven Protein Data Bank, April 1993
A:Reference number: A51894; PDB:1OM6
A:Contents: annotation, confirmation by ¹H NMR, residues 46-72
R:Davis, J.H.; Bradley, F.K.; Miljanich, G.P.; Nadashi, I.; Ramchandran, J.; Basus, Biochemistry 32, 7392-7405, 1993
A:Title: Solution structure of omega-conotoxin GVIA using 2-D NMR spectroscopy and Fourier transform
A:Reference number: A58326; MBI:94332945
A:Contents: annotation, confirmation by ¹H NMR

Ref: Lohman, B.M.; Pennington, M.W.; Norton, R.S.
Submitted to the Brookhaven Protein Data Bank, August 1993

Accession number: A51089; PDB: 10X

Abstract: Structural conformation by (1)H-NMR, residues 46-72

Comment: There are several types of conotoxins: alpha, acting on postsynaptic membrane

neurotoxic.

Keywords: alpha-conotoxin
Keywords: acetylcholine release inhibition; amidated carboxyl end; calcium channel inh

11 22/10/93: status predicted <SIG>

12 24/4/94: status predicted <PRO>

13 24/7/94: status predicted <MAT>

14 24/7/94: status predicted <MAT>

15 24/7/94: status predicted <MAT>

16 24/7/94: status predicted <MAT>

17 24/7/94: status predicted <MAT>

18 24/7/94: status predicted <MAT>

19 24/7/94: status predicted <MAT>

20 24/7/94: status predicted <MAT>

21 24/7/94: status predicted <MAT>

22 24/7/94: status predicted <MAT>

23 24/7/94: status predicted <MAT>

24 24/7/94: status predicted <MAT>

25 24/7/94: status predicted <MAT>

26 24/7/94: status predicted <MAT>

27 24/7/94: status predicted <MAT>

28 24/7/94: status predicted <MAT>

29 24/7/94: status predicted <MAT>

30 24/7/94: status predicted <MAT>

31 24/7/94: status predicted <MAT>

32 24/7/94: status predicted <MAT>

33 24/7/94: status predicted <MAT>

34 24/7/94: status predicted <MAT>

35 24/7/94: status predicted <MAT>

36 24/7/94: status predicted <MAT>

37 24/7/94: status predicted <MAT>

38 24/7/94: status predicted <MAT>

39 24/7/94: status predicted <MAT>

40 24/7/94: status predicted <MAT>

41 24/7/94: status predicted <MAT>

42 24/7/94: status predicted <MAT>

43 24/7/94: status predicted <MAT>

44 24/7/94: status predicted <MAT>

45 24/7/94: status predicted <MAT>

46 24/7/94: status predicted <MAT>

47 24/7/94: status predicted <MAT>

48 24/7/94: status predicted <MAT>

49 24/7/94: status predicted <MAT>

50 24/7/94: status predicted <MAT>

51 24/7/94: status predicted <MAT>

52 24/7/94: status predicted <MAT>

53 24/7/94: status predicted <MAT>

54 24/7/94: status predicted <MAT>

55 24/7/94: status predicted <MAT>

56 24/7/94: status predicted <MAT>

57 24/7/94: status predicted <MAT>

58 24/7/94: status predicted <MAT>

59 24/7/94: status predicted <MAT>

60 24/7/94: status predicted <MAT>

61 24/7/94: status predicted <MAT>

62 24/7/94: status predicted <MAT>

63 24/7/94: status predicted <MAT>

64 24/7/94: status predicted <MAT>

65 24/7/94: status predicted <MAT>

66 24/7/94: status predicted <MAT>

Search completed: April 19, 2001, 06:34:04

Job time: 549 sec

1

2

3

Future version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

M. Fortin - prof. in Sociol., using 50 models

Run on: April 19, 2001, 10:54:24 Search time 95.86 Seconds
(without alignments)
55.021 Million cell updates/second

```

141101  DS-08-906-465-Z_00PY_1_45
141102  Foot Foot Score: 241
141103  Sequences: 1 M01KNS221M222M223M224M225M226M227M228M229M230M231M232M233M234M235M236M237M238M239M240M241M242M243M244M245M246M247M248M249M250M251M252M253M254M255M256M257M258M259M260M261M262M263M264M265M266M267M268M269M270M271M272M273M274M275M276M277M278M279M280M281M282M283M284M285M286M287M288M289M290M291M292M293M294M295M296M297M298M299M300M301M302M303M304M305M306M307M308M309M310M311M312M313M314M315M316M317M318M319M320M321M322M323M324M325M326M327M328M329M330M331M332M333M334M335M336M337M338M339M340M341M342M343M344M345M346M347M348M349M350M351M352M353M354M355M356M357M358M359M360M361M362M363M364M365M366M367M368M369M370M371M372M373M374M375M376M377M378M379M380M381M382M383M384M385M386M387M388M389M390M391M392M393M394M395M396M397M398M399M400M401M402M403M404M405M406M407M408M409M410M411M412M413M414M415M416M417M418M419M420M421M422M423M424M425M426M427M428M429M430M431M432M433M434M435M436M437M438M439M440M441M442M443M444M445M446M447M448M449M450M451M452M453M454M455M456M457M458M459M460M461M462M463M464M465M466M467M468M469M470M471M472M473M474M475M476M477M478M479M480M481M482M483M484M485M486M487M488M489M490M491M492M493M494M495M496M497M498M499M500M501M502M503M504M505M506M507M508M509M510M511M512M513M514M515M516M517M518M519M520M521M522M523M524M525M526M527M528M529M530M531M532M533M534M535M536M537M538M539M540M541M542M543M544M545M546M547M548M549M550M551M552M553M554M555M556M557M558M559M560M561M562M563M564M565M566M567M568M569M570M571M572M573M574M575M576M577M578M579M580M581M582M583M584M585M586M587M588M589M590M591M592M593M594M595M596M597M598M599M600M601M602M603M604M605M606M607M608M609M610M611M612M613M614M615M616M617M618M619M620M621M622M623M624M625M626M627M628M629M630M631M632M633M634M635M636M637M638M639M640M641M642M643M644M645M646M647M648M649M650M651M652M653M654M655M656M657M658M659M660M661M662M663M664M665M666M667M668M669M670M671M672M673M674M675M676M677M678M679M680M681M682M683M684M685M686M687M688M689M690M691M692M693M694M695M696M697M698M699M700M701M702M703M704M705M706M707M708M709M710M711M712M713M714M715M716M717M718M719M720M721M722M723M724M725M726M727M728M729M730M731M732M733M734M735M736M737M738M739M740M741M742M743M744M745M746M747M748M749M750M751M752M753M754M755M756M757M758M759M760M761M762M763M764M765M766M767M768M769M770M771M772M773M774M775M776M777M778M779M780M781M782M783M784M785M786M787M788M789M790M791M792M793M794M795M796M797M798M799M800M801M802M803M804M805M806M807M808M809M810M811M812M813M814M815M816M817M818M819M820M821M822M823M824M825M826M827M828M829M830M831M832M833M834M835M836M837M838M839M840M841M842M843M844M845M846M847M848M849M850M851M852M853M854M855M856M857M858M859M860M861M862M863M864M865M866M867M868M869M870M871M872M873M874M875M876M877M878M879M880M881M882M883M884M885M886M887M888M889M890M891M892M893M894M895M896M897M898M899M900M901M902M903M904M905M906M907M908M909M910M911M912M913M914M915M916M917M918M919M920M921M922M923M924M925M926M927M928M929M930M931M932M933M934M935M936M937M938M939M940M941M942M943M944M945M946M947M948M949M950M951M952M953M954M955M956M957M958M959M960M961M962M963M964M965M966M967M968M969M970M971M972M973M974M975M976M977M978M979M980M981M982M983M984M985M986M987M988M989M990M991M992M993M994M995M996M997M998M999

```

Copyright 1991, Oxford University Press

Number of hits satisfying chosen parameters: 374700

Maximum Likelihood Estimation

Final Position: Minimum Match

Listed first 4, "Linear"ies

```

1: SP_kernel_150.*
2: SP_barotra.*
3: SP_bird.*
4: SP_birdm.*
5: SP_invertebrate.*
6: SP_mammal.*
7: SP_mice.*
8: SP_orchardlot.*
9: SP_phages.*
10: SP_plant.*
11: SP_fishoil.*
12: SP_virus.*
13: SP_zootera.*
14: SP_unclassified.*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARY

Result	Score	Quality	Match	Length	DB	ID	Description
1	241	100.0	69	4	090803	090803	homo sapien
2	241	100.0	42	4	090809	090809	homo sapien
3	237	90.0	50	4	095171	095171	caallitrix
4	208	84.2	139	11	093097	093097	mus musculus
5	198	60.1	163	6	09XSW2	09XSW2	sus scrofa
6	192	60.1	127	4	09X3C5	09X3C5	bos taurus
7	193	80.1	527	6	091015	091015	ovis aries
8	141	58.5	13	4	090VE2	090VE2	sturnus
9	73	40.3	473	13	014012	014012	aquila
10	66	27.3	421	5	09Y1X5	09Y1X5	ephausa
11	62.5	25.9	481	2	066144	066144	citrobacter
12	62	25.7	396	6	028874	028874	canis
13	61.5	25.5	548	13	098669	098669	carassius
14	57	24.5	209	10	064873	064873	arabidopsi
15	59	24.5	390	5	091193	091193	caenorhabdi
16	57.5	24.9	217	5	090VX3	090VX3	drosofila
17	57	23.7	327	5	025688	025688	plasmodium
18	57	23.7	396	6	094717	094717	schizosaccha
19	56.5	23.3	330	5	021423	021423	schizosaccha

20	56.5	22.4	7.97	5	QVAV1	QVAV1	droscophila
21	55	22.1	2.47	2	QVAB7.2	QVAB7.2	arabidopsis
22	56	23.2	55.3	4	QVAB8	QVAB8	brassica
23	56	23.2	9.07	4	QVAB2.7	QVAB2.7	brassica
24	55.5	23.2	4823	1	QVAB2.1	QVAB2.1	brassica
25	55.5	23.0	3.32	1	QVAB7.5	QVAB7.5	arabidopsis
26	55.5	23.0	5.91	11	QVAB7.9	QVAB7.9	arabidopsis
27	55	22.8	1.04	10	QVAB8	QVAB8	arabidopsis
28	55	22.8	1785	5	QVAB8.5	QVAB8.5	arabidopsis
29	55	22.8	1329	5	QVAB7.7	QVAB7.7	arabidopsis
30	54.5	22.5	330	4	QVAB7.2	QVAB7.2	arabidopsis
31	54.5	22.6	424	5	QVAB8.8	QVAB8.8	arabidopsis
32	54.5	22.6	561	8	QVAB17	QVAB17	arabidopsis
33	54.5	22.6	75.9	12	QVAB12	QVAB12	arabidopsis
34	54.5	22.6	869	5	QVAB14	QVAB14	arabidopsis
35	54.5	22.6	928	10	QVAB15	QVAB15	arabidopsis
36	54.5	22.6	968	5	QVAB9	QVAB9	arabidopsis
37	54.5	22.6	977	3	QVAB5	QVAB5	arabidopsis
38	54.5	22.6	1036	3	QVAB7.8	QVAB7.8	arabidopsis
39	54.5	22.6	1231	5	QVAB2.4	QVAB2.4	arabidopsis
40	54	22.4	377	2	QVAB7.1	QVAB7.1	arabidopsis
41	54	22.4	474	3	QVAB7.1	QVAB7.1	arabidopsis
42	54	22.4	556	11	QVAB5	QVAB5	arabidopsis
43	54	22.4	611	7	QVAB4	QVAB4	arabidopsis
44	54	22.4	1011	4	QVAB7.3	QVAB7.3	arabidopsis
45	54	22.4	1131	4	QVAB8	QVAB8	arabidopsis

ANNOUNCEMENT

RESULT	1
G00H03	
ID	G00H03
AC	G00H03
D1	01-MAY-2000 (ITEMB.ref), 13, (treated)
D1	01-MAY-2000 (ITEMB.ref), 13, last sequence update)
D1	01-MAY-2000 (ITEMB.ref), 13, last annotation update)
DE	ESTROGEN RECEPTOR BETA (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Primates; Catartida; Hominoidea; Homin-
OX	NHLtaxid 96067.
RN	[1]
RP	SEQUENCE FROM N.A.
RA	L.I.C., Danbury, Ct.
PT	"Cloning and characterization of the estrogen receptor beta gene
RT	promoter."
RL	Submitted (87-1965) to the EMBL/Genbank/Jdb database.
DR	EMBL: AF191443; AAF2432.1.
KW	Receptor.
FT	NON TER
SQ	SEQUENCE 69 AA, 7825 MW; 4159 Amino Acids; 160 kDa.
<hr/>	
Query Match	100.0% Score 241.76 bits Identical 97%
Best Local Similarity	100.0% Ident No. 20/24
Matches	452 conservative 0% Miscellaneous 12 Indels 62 Mismatches
<hr/>	
1	MEFNSTSLSSLSSTNSLSTSLTCTGCTGGTCTGGTAAAT
1	MDKHSSTSLSPSYNSVNSTGLFNSVLDSGVSHHHYTAAT
<hr/>	
RESULT	2
G00U09	
ID	G00U09
AC	G00U09
D1	01-MAY-2000 (ITEMB.ref), 13, (treated)
D1	01-MAY-2000 (ITEMB.ref), 13, last sequence update)
D1	01-MAY-2000 (ITEMB.ref), 13, last annotation update)
DE	ESTROGEN RECEPTOR BETA2 SPALTY VARIANT.
OS	Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID:9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 KA Lo H., Leygue E., Dorelas H., Watson P.H., Murphy L.C.;
 KT Cloning and functional characteristics of human estrogen receptor
 beta isoforms.
 DT Submitted (JAN 1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF124790; ADD3580.1;
 DR HSSP: P03372; 1HCP.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR001628;
 DR PFAM: PF00105; 21-c4.1.
 DE PRINTS: PR00647; STROGDRNDR.
 DR PRINTS: PR00350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
 KW Receptor.
 SQ SEQUENCE 323 AA: 35944 MW: 158376056D3CA12 CRC64;

Query Match 100.0%; Score 241; DB 4; Length 323;
 Best Local Similarity 100.0%; Pred. No. 1c-23;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MDKNSPSSLNPSSTNCSSQSLPLEHGSIVIPSSVDSHHEVPA 45
 DB 1 MDKNSPSSLNPSSTNCSSQSLPLEHGSIVIPSSVDSHHEVPA 45

RESULT 3
 OY5171 PRELIMINARY: PRT: 530 AA.
 AC OY5171;
 DT 01-FEB-1997 (TREMBLER, 02, Created)
 DT 01-MAY-2000 (TREMBLER, 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLER, 15, Last annotation update)
 DE ESTROGEN RECEPTOR BETA.
 OS Callithrix jacchus (Common marmoset).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callitrix.
 OX NCBI_TaxID:9483;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 KA Gaughan J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS.
 KA Noble G.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC 1-1 SCHMIDTKE LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: Y09472; CAAT0546.2;
 DR HSSP: P03372; 1HCP.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR005236;
 DR INTERPRO: IPR001628;
 DR INTERPRO: IPR001723;
 DR PFAM: PF00104; hormone_rec_1.
 DR PFAM: PF00105; 21-c4.1.
 DE PRINTS: PR00047; STROGDRNDR.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00398; STROGDRNDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger.
 SQ SEQUENCE 530 AA: 59087 MW: 45089107A44C5301 CRC64;

Query Match 90.0%; Score 217; DB 6; Length 530;
 Best Local Similarity 91.1%; Pred. No. 2.4e-20;

Matches 41; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 1 MDKNSPSSLNPSSTNCSSQSLPLEHGSIVIPSSVDSHHEVPA 45
 DB 1 MDKNSPSSLNPSSTNCSSQSLPLEHGSIVIPSSVDSHHEVPA 45

RESULT 4
 OYJ9M97 PRELIMINARY: PRT: 139 AA.
 ID OYJ9M97;
 DT 01-OCT-2000 (TREMBLER, 15, Created)
 DT 01-OCT-2000 (TREMBLER, 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLER, 15, Last annotation update)
 DE ESTROGEN RECEPTOR BETA (FRAGMENT).
 CN ERB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID:10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 KA Ishihashi O., Kawashima H.;
 KT "Cloning and characterization of the promoter region of mouse estrogen
 receptor beta."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB044957; BAA92313.1;
 KW Receptor.
 SQ SEQUENCE 139 AA: 15215 MW: 4703615A8E5940AB CRC64;

Query Match 84.2%; Score 203; DB 11; Length 139;
 Best Local Similarity 84.4%; Pred. No. 3.9e-19;
 Matches 39; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 MDKNSPSSLNPSSTNCSSQSLPLEHGSIVIPSSVDSHHEVPA 45
 DB 20 MDKNSPSSLNPSSTNCSSQSLPLEHGSIVIPSSVDSHHEVPA 64
 RESULT 5
 OYXSW2 PRELIMINARY: PRT: 163 AA.
 ID OYXSW2;
 DT 01-NOV-1999 (TREMBLER, 12, Created)
 DT 01-NOV-1999 (TREMBLER, 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLER, 15, Last annotation update)
 DE ESTROGEN RECEPTOR BETA (FRAGMENT).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae Sus.
 OX NCBI_TaxID:9623;
 RN 11
 RP SEQUENCE FROM N.A.
 KA Kowalski A.A., Grady L.S., Stamen E.C.M.;
 RL "Expression of estrogen receptor beta by the porcine embryo during the
 preimplantation period and its possible role in the regulation of
 growth."
 DT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF164957; AAD45381.1;
 DR HSSP: P03372; 1HCP.
 DR INTERPRO: IPR001628;
 DR PFAM: PF00105; 21-c4.1.
 KW Receptor.
 SQ SEQUENCE 163 AA: 17704 MW: 449641A3B63181E CRC64;

Query Match 80.1%; Score 193; DB 6; Length 163;
 Best Local Similarity 82.2%; Pred. No. 9.4e-18;
 Matches 37; Conservative 3; Mismatches 5; Indels 0; Gaps 0;


```

013012
ID 013012 PRELIMINARY: PRT: 573 AA
AC 013012
DT 01-JUL-1997 (TREMblrel: 04, created)
D1 01-JUL-1997 (TREMblrel: 04, last sequence update)
D2 01-OCT-2000 (TREMblrel: 15, last annotation update)
DE ESTROGEN RECEPTOR.
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OX Anguillidae; Anguilla.
PN 11
RC TISSUE: LIVER.
RF SEQUENCE FROM N.A.
RA Tada T.;
RL Submitted (AFR-1997) to the EMBL/Genbank/DBP1 databases.
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE: LIVER.
RA MEDLINE: 96396022; PubMed: 8793852;
RA Tada T.; Adachi S.; Yamauchi K.;
RT "Molecular cloning and characterization of Japanese eel estrogen
receptor cDNA."
RL Mol. Cell. Endocrinol. 119:37-45(1996).
CC 1- SURCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DE EMBL: AB003356; HAA19651.1; -.
DR HSSP: P03372; IMGT.
DR INTERPRO: IPR000536; -.
DR INTERPRO: IPR001628; -.
DR INTERPRO: IPR001723; -.
DR PFAM: PF00104; hormone_rec_1.
DR PFAM: PF00105; zf-C4; 1.
DR PRINTS: PR00047; STROLDPFRNGR.
DR PRINTS: PR00198; STROLDPFRNGR.
DR PROSITE: PS00033; NUCLEAR_PEPFPRK; 1
DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
SO SEQUENCE 573 AA; 63420 MW; 9C64C1D8D39ED4CC CRC64;

Query Match 30 34: Score 73. DB 13: Length 573;
Best Local Similarity 42.48; Pred. No. 0.17;
Matches 14: Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 1 PSSYNCSSSLIPLEHGSIVIPSSVYSHHEVPA 45
: : : : : : : : : : : : : : : :
DB 33 PIMYKALPALSMESHAVGIPSPYDSSHDYAA 65

RESULT 10
OYV1X6 PRELIMINARY: PRT: 471 AA.
AC 0YV1X6
DT 01-NOV-1999 (TREMblrel: 12, created)
D1 01-NOV-1999 (TREMblrel: 12, last sequence update)
D2 01-OCT-2000 (TREMblrel: 15, last annotation update)
DE SPTR4 (EC 3.1.3.48) (FRAGMENT).
OS Ephydriidae; Ephydriidae.
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceraclinozoa;
OC Haptosclerida; Spongillidae; Ephydria.
OX NCBI_TaxID: 31330;
PN 11
RP SEQUENCE FROM N.A.
RA MEDLINE: 94246376; PubMed: 10229569;
RA Ono K.; Suda H.; Iwabe N.; Kuma K.; Miyata T.;
RT "Multiple protein tyrosine phosphatases in sponges and explosive gene
duplication in the early evolution of animals before the palaeozoan-
erian eozoan split."
RL J. Mol. Evol. 48:654-662(1999).
DR EMBL: AB019125; HAA82558.1; -.
DR HSSP: P18052; TYRO.
DR INTERPRO: IPR000242; -.

```

```

DR INTERPRO: IPR000387; -.
DR PFAM: PF00102; Y_phosphatase_2.
DR PRINTS: PR00100; PKTIPPHTRASE.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
DR PROSITE: PS00055; TYR_PHOSPHATASE_PTP; 2.
DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 2.
KW Hydrolyase.
FT NON-TER.
SO SEQUENCE 471 AA; 54509 MW; 7DA62456154A6A56 CRC64;

Query Match 27.48; Score 66; DB 5: Length 471;
Best Local Similarity 32.56; Pred. No. 1.1;
Matches 13: Conservative 11; Mismatches 16; Indels 0; Gaps 0;

OY 3 IKNSPSSLSNPSYNCSSSLIPLEHGSIVIPSSVYSHHEVPA 42
: : : : : : : : : : : : : : : :
DB 227 VERKINIVFUSHHETMSLIFGVESSTIMASLTCYHKE 266

RESULT 11
OYV144 PRELIMINARY: PRT: 481 AA.
AC 0YV144
DT 01-AUG-1998 (TREMblrel: 07, created)
D1 01-AUG-1998 (TREMblrel: 07, last sequence update)
D2 01-AUG-1998 (TREMblrel: 07, last annotation update)
DE GI07AMATE M07ASF.
OS Citrobacter amaloncus.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Citrobacter.
OX NCBI_TaxID: 35703;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN: YG-1002;
RA Kato Y.; Asano Y.;
RL Appl. Microbiol. Biotechnol. 0:0-0(1998).
DR EMBL: AB005294; BAA28708.1; -.
SO SEQUENCE 481 AA; 52942 MW; DEFD7E51E6647751 CRC64;

Query Match 25.98; Score 62.5; DB 2: Length 481;
Best Local Similarity 34.78; Pred. No. 3.3;
Matches 17: Conservative 10; Mismatches 17; Indels 5; Gaps 2;

OY 1 MDIKNSPSSLSNPSYNCSSSLIPLEHGSIVIPSSVYSHHEVPA 45
: : : : : : : : : : : : : : : :
DB 403 LDYFAPASGNACKILPVKRN SGALIKVLPAGSVPLKDLALHHDYVA 450

RESULT 12
OYV144 PRELIMINARY: PRT: 496 AA.
AC 0YV144
DT 01-NOV-1996 (TREMblrel: 01, created)
D1 01-NOV-1996 (TREMblrel: 01, last sequence update)
D2 01-JUL-2000 (TREMblrel: 14, last annotation update)
DE TISSUE FACTOR PATHWAY INHIBITOR.
OS Canis familiaris (dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID: 9615;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE: 95071310; PubMed: 7980463;
RA Girard T.J.; Gallant D.; Broze G.J., Jr.;
RT "Complementary DNA sequencing of canine tissue factor pathway
inhibitor reveals a unique nanomeric repetitive sequence between the
second and third Knitz domains."
RL Biochem. J. 303:923-928(1994).
DR EMBL: S73369; AAH32443.1; -.
DR HSSP: P10646; ITFX.
DR INTERPRO: IPR02223; -.

```

Methodology: endophytes; core endophytes; Rosidae; eucosids. II

[illegible]

Fri Apr 20 09:01:00 2001

us-08-906-365-2_copy_1_45.rspt

Page 6

DB 78 LKOSRPSPLSNSTPSPVPIQNLJEDRTPKLSPELPSERGNISFTLST-DSH 133

Search completed: April 19, 2001, 10:34:28
Job time: 14736 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

ORF protein : protein search, using SW model.

Run on: April 19, 2001, 10:45:31 : Search time 51.46 Seconds

(without alignments)

29,955 Million cell updates/sec

Hit list: US-08-906-365-2_copy_1_45

Sequence: 1 MURKSPSSINSPSSVYNQ.....EHSIVPSVYSHHEVYA 45

Scoring Table: HUSUM62

Gapop 10.0, Gapext 0.5

Searched: 94435 seqs, 44255486 residues

Total number of hits satisfying chosen parameters: 94435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 10%

Listing first 45 summaries

Database: 1 SwissProt_1994*

Note: Results are the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	IO	Description
1	241	100.0	530	1	ESR2_HUMAN
2	219	90.0	530	1	ESR2_VALIA
3	205	85.1	540	1	ESR2_RAT
4	203	84.2	540	1	ESR2_MOUSE
5	193	80.1	164	1	ESR2_PIG
6	193	80.1	527	1	ESR2_BOVIN
7	193	80.1	527	1	ESR2_SHEP
8	141	58.5	554	1	ESR2_STUV
9	73	30.3	573	1	ESR2_ANCIA
10	61.5	25.5	508	1	ERH1_GARAU
11	58.5	24.3	664	1	YFEL_YEAST
12	57	23.7	598	1	AMAL_PLANT
13	56.3	23.4	612	1	YRBB_YEAST
14	56	23.2	273	1	YFHA_YEAST
15	56	23.2	445	1	HUSX_HUMAN
16	56	23.2	743	1	ILF2_HUMAN
17	56	23.2	1315	1	ALC9_SCHU
18	55.5	23.0	354	1	AMPK_MYCN
19	55.5	23.0	673	1	ESR2_MICR
20	55	22.8	73	1	FOXO_GONOR
21	55	22.8	243	1	G112_RAT
22	55	22.8	766	1	EYA2_HUMAN
23	55	22.8	909	1	AM12_HUMAN
24	54.5	22.5	335	1	D4H1_RAT
25	54.5	22.5	868	1	SRVC_DROME
26	54	22.1	213	1	G112_MOUSE
27	54	22.4	338	1	Y4X1_YEAST
28	53.5	22.2	612	1	YND4_YEAST
29	53.5	22.2	1146	1	YH32_YEAST
30	53	22.0	619	1	M4K2_MOUSE
31	53	22.0	1187	1	ETRE_HUMAN
32	52.5	21.8	449	1	M553_NYAC
33	52.5	21.8	1176	1	CTAA_BACIK

34	52.5	21.8	1176	1	CTAA_BACIK
35	52.5	21.8	1176	1	CTAA_BACIK
36	52	21.6	107	1	FER_PSAIA
37	52	21.6	559	1	ESR2_PSAIA
38	52	21.6	636	1	YRBB_YEAST
39	52	21.6	647	1	YRBB_YEAST
40	52	21.6	1014	1	MUL_HUMAN
41	52	21.6	1189	1	CTAA_BACIK
42	51.5	21.4	346	1	CTAA_HUMAN
43	51.5	21.4	349	1	YH28_PIG
44	51.5	21.4	575	1	ESR2_PIG
45	51	21.2	246	1	YH28_PIG

ALIGNMENTS

RESULT 1
ID ESR2_HUMAN STANDBY: PRI: 530 AA
AC G927317 (006653) (006087) (075584) (007023) (007023) (075584) (006653)
DT 15-JUL-1999 (Ref. 38) (Created)
DI 15-JUL-1999 (Ref. 38) (Last sequence update)
DE 01-OCT-2000 (Ref. 45) (Last annotation update)
DE ESTROGEN RECEPTOR BETA (ER-BETA).
CN ESR2 OR NR3A2 OR ESTRB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Theria; Primates; Eularchia; Hominoidea; Homo.
OX NCBI_Taxid 9606;
RN (1)
RP SEQUENCE FROM N.A. (USP-BM BETA-1).
RX MEDLINE=98139678; PubMed=947341;
RA Ogawa S., Inoue S., Watanabe T., Hiroi H., Ogawa A., Hasegawa T.,
Oguchi Y., Muramatsu M.;
RT "The complete primary structure of human estrogen receptor beta
(ER-beta) and its heterodimerization with ER-alpha in vivo and in
vitro.";
RI Biochem. Biophys. Res. Commun. 241:122-129 (1998).
RI (2)
RN SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RP TISSUE-TOSTIS;
RX MEDLINE=98404286; PubMed=96667;
RA Moselman S., Polman J., Dijkema P.;
RT "ER beta: identification and characterization of a novel human
estrogen receptor.";
RI FEBS Lett. 492:44-50 (1996).
RN (3)
RN SEQUENCE OF 46 530 FROM N.A. (USP-BM BETA 1), AND CHARACTERIZATION
OF TISSUE-TOSTIS;
RX MEDLINE=96454675; PubMed=879613;
RA Moselman S., Polman J., Dijkema P.;
RT "ER beta: identification and characterization of a novel human
estrogen receptor.";
RI FEBS Lett. 492:44-50 (1996).
RN (4)
RN SEQUENCE FROM N.A. (USP-BM BETA 2), AND CHARACTERIZATION
OF TISSUE-TOSTIS;
RX MEDLINE=98437908; PubMed=971411;
RA Ogawa S., Inoue S., Watanabe T., Ogawa A., Hiroi H.,
Muramatsu M.;
RT "Molecular cloning and characterization of human estrogen receptor
beta (a potential inhibitor of estrogen action in human).";
RI Nucleic Acids Res. 27:505-512 (1999).
RN (5)
RN SEQUENCE OF 54 530 FROM N.A. (USP-BM BETA 1A AND BETA 1);
RP TISSUE-Endometrium;
RX Brandenberger A.W., Lefebvre D., Layan B.N., Lefebvre P.M.;
RT "Subcloning and sequencing of the ER-beta gene from human
endometrium.";
RI Sequence from N.A. (USP-BM BETA 2A).


```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announe/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: AF110402; AAD24432.1; -.
DR EMBL: Y18017; CAH53861.1; ALT_INIT.
DR HSSP: P03572; IHC1.
DR InterPro: IPR000524; -.
DR InterPro: IPR000536; -.
DR InterPro: IPR001628; -.
DR InterPro: IPR001723; -.
DR InterPro: IPR001728; -.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zf-C4_1.
DR PRINTS: PR00047; STEROIDINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PRINTS: PR00398; STEROIDOMER.
DR PRINTS: PR00546; THYROIDOMER.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Phosphorylation.
KW Zinc-finger; Steroid-binding; Modulating.
FT DOMAIN 1 145
FT DNA_BIND 146 211 NUCLEAR_RECEPTOR-TYPE.
FT ZN_FING 146 166 C4-TYPE.
FT ZN_FING 182 206 C4-TYPE.
FT DOMAIN 212 527 STEROID-BINDING (BY SIMILARITY).
FT MOD_RES 84 84 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 485 485 PHOSPHORYLATION (BY SIMILARITY).
FT CONFLICT 60 60 N > D (IN REF. 2).
SQ SEQUENCE 527 AA; 5601 MW; 9CFPF06F4FC64 CRC64;

Query Match 80.1%; Score 193; DB 1; Length 527;
Best Local Similarity 80.0%; Pred. No. 9,6e-17;
Matches 36; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDKNSPSSLSNPSVYNGQSILPLEEGSYIPSSVYSDSHHEPA 45
DB 1 MDKNSPSSLSNPSVYNGQSILPLEEGSYIPSSVYSDSHHEPA 45

RESULT 7
ESR2_SHEEP STANDARD; PRT: 527 AA.
AC Q9YU15; Q9N0T6;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ESTROGEN RECEPTOR BETA (ER-BETA).
GN ESR2 OR NR3A2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eulambda; Cetartiodactyla; Ruminantia; Bovidae; Bovinae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A. (ISOPFORMS BETA AND BETA-1).
KC TISSUE=Ovary;
RA Cardenas H., Bissby K.M., Pope W.F., Nephew K.P.;
RT "Cloning and expression of estrogen receptor beta in the ovine
RT ovary";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
CC ER-ALPHA. AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
CC ALPHA (BY SIMILARITY).
CC -1- SIMILAR LIGAND LOCATIONS: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: BETA (SHOWN HERE) AND
CC BETA-1; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.

```

```

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation in
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announe/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: AF177836; AAD55772.1; -.
DR EMBL: AF257109; AAF71745.1; -.
DR HSSP: P03572; IHC1.
DR InterPro: IPR000324; -.
DR InterPro: IPR000536; -.
DR InterPro: IPR001628; -.
DR InterPro: IPR001723; -.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zf-C4_1.
DR PRINTS: PR00047; STEROIDINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PRINTS: PR00398; STEROIDOMER.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor: Transcription regulation; DNA binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation.
KW Zinc-finger; Steroid-binding; Modulating.
FT DOMAIN 1 145
FT DNA_BIND 146 211 NUCLEAR_RECEPTOR-TYPE.
FT ZN_FING 146 166 C4-TYPE.
FT ZN_FING 182 206 C4-TYPE.
FT DOMAIN 212 527 STEROID-BINDING (BY SIMILARITY).
FT MOD_RES 84 84 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 485 485 PHOSPHORYLATION (BY SIMILARITY).
FT VARSPLIC 316 324 FVLSLYSDQ -> MKSNVLEE (IN ISOFORM BETA
FT VARSPLIC 1).
SQ SEQUENCE 527 AA; 59120 MW; 90D7A3B0947E1EA CRC64;

Query Match 80.1%; Score 193; DB 1; Length 527;
Best Local Similarity 80.0%; Pred. No. 9,6e-17;
Matches 36; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDKNSPSSLSNPSVYNGQSILPLEEGSYIPSSVYSDSHHEPA 45
DB 1 MDKNSPSSLSNPSVYNGQSILPLEEGSYIPSSVYSDSHHEPA 45

RESULT 8
ESR2_STUVU STANDARD; PRT: 554 AA.
AC Q9PVE2;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ESTROGEN RECEPTOR BETA (ER-BETA).
GN ESR2 OR NR3A2.
OS Styrax vulgaris (Starling).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Cyprinodontiformes; Styraxidae; Styrax.
OX NCBI_TaxID=9172;
RN [1]
RP SEQUENCE FROM N.A.
KC TISSUE=Brain;
RX MEDLINE=94427866; PubMed=10490520;
RA Bernard D.J., Bentley G.E., Battiazzi J., Turck F.M., Ball G.P.;
RT "A human estrogen receptor alpha and estrogen receptor
RT beta show distinct patterns of expression in rodent and
RT nuclei of European starlings";
RL Endocrinology 140:4633-4643(1999).
CC -1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF ER-
CC ALPHA. AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN DEPENDENT MANNER.

```

```

00 1. LOCALLY SYNTHESIZED ESTROGENS MAY ACT VIA ER ALPHA, IN ADDITION TO
01 ER ALPHA, TO MODULATE SEASONAL OR DEVELOPMENTAL EFFECTS ON NEARBY
02 SEX-NOBLE.
03 1. SUBUNIT BINDS DNA AS A HOMO-DIMER. CAN FORM A HETERODIMER WITH ER-
04 ALPHA (BY SIMILARITY).
05 1. SUBCELLULAR LOCATION: NUCLEAR.
06 1. TISSUE SPECIFICITY: BRAIN, PITUITARY, SKELETAL MUSCLE, LIVER,
07 ADIPONAL, KIDNEY, INTESTINE, AND OVARY.
08 1. DOMAIN: COMPOSED OF THREE DOMAINS, A MODULATING N-TERMINAL DOMAIN,
09 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
10 1. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
11 NR3 SUPERFAMILY.
12
13 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
14 between the Swiss Institute of Bioinformatics and the EMBL outstation
15 at the European Bioinformatics Institute. There are no restrictions on its
16 use by non-profit institutions as long as its content is in no way
17 modified and this statement is not removed, usage by and for commercial
18 entities requires a license agreement (see http://www.ebi.ac.uk/Ensembl/
19 or send an email to license@ebi.ac.uk).
20
21 EMBL: AF114513; AAC056993.1;
22 HSSP: P04372; H0CP.
23 InterPro: IPR000824;
24 InterPro: IPR001628;
25 InterPro: IPR001723;
26 InterPro: IPR001729;
27 PRINIS: PR000477; STROIDRECEPTOR;
28 PRINIS: PR00450; VITAMINER;
29 PRINIS: PR00498; STEROIDRECEPTOR;
30 PRINIS: PR00498; STEROIDRECEPTOR;
31 PRINIS: PR00498; STEROIDRECEPTOR;
32 PRINIS: PR00498; STEROIDRECEPTOR;
33 PRINIS: PR00498; STEROIDRECEPTOR;
34 PRINIS: PR00498; STEROIDRECEPTOR;
35 PRINIS: PR00498; STEROIDRECEPTOR;
36 PRINIS: PR00498; STEROIDRECEPTOR;
37 PRINIS: PR00498; STEROIDRECEPTOR;
38 PRINIS: PR00498; STEROIDRECEPTOR;
39 PRINIS: PR00498; STEROIDRECEPTOR;
40 PRINIS: PR00498; STEROIDRECEPTOR;
41 PRINIS: PR00498; STEROIDRECEPTOR;
42 PRINIS: PR00498; STEROIDRECEPTOR;
43 PRINIS: PR00498; STEROIDRECEPTOR;
44 PRINIS: PR00498; STEROIDRECEPTOR;
45 PRINIS: PR00498; STEROIDRECEPTOR;
46 PRINIS: PR00498; STEROIDRECEPTOR;
47 PRINIS: PR00498; STEROIDRECEPTOR;
48 PRINIS: PR00498; STEROIDRECEPTOR;
49 PRINIS: PR00498; STEROIDRECEPTOR;
50 PRINIS: PR00498; STEROIDRECEPTOR;
51 PRINIS: PR00498; STEROIDRECEPTOR;
52 PRINIS: PR00498; STEROIDRECEPTOR;
53 PRINIS: PR00498; STEROIDRECEPTOR;
54 PRINIS: PR00498; STEROIDRECEPTOR;
55 PRINIS: PR00498; STEROIDRECEPTOR;
56 PRINIS: PR00498; STEROIDRECEPTOR;
57 PRINIS: PR00498; STEROIDRECEPTOR;
58 PRINIS: PR00498; STEROIDRECEPTOR;
59 PRINIS: PR00498; STEROIDRECEPTOR;
60 PRINIS: PR00498; STEROIDRECEPTOR;
61 PRINIS: PR00498; STEROIDRECEPTOR;
62 PRINIS: PR00498; STEROIDRECEPTOR;
63 PRINIS: PR00498; STEROIDRECEPTOR;
64 PRINIS: PR00498; STEROIDRECEPTOR;
65 PRINIS: PR00498; STEROIDRECEPTOR;
66 PRINIS: PR00498; STEROIDRECEPTOR;
67 PRINIS: PR00498; STEROIDRECEPTOR;
68 PRINIS: PR00498; STEROIDRECEPTOR;
69 PRINIS: PR00498; STEROIDRECEPTOR;
70 PRINIS: PR00498; STEROIDRECEPTOR;
71 PRINIS: PR00498; STEROIDRECEPTOR;
72 PRINIS: PR00498; STEROIDRECEPTOR;
73 PRINIS: PR00498; STEROIDRECEPTOR;
74 PRINIS: PR00498; STEROIDRECEPTOR;
75 PRINIS: PR00498; STEROIDRECEPTOR;
76 PRINIS: PR00498; STEROIDRECEPTOR;
77 PRINIS: PR00498; STEROIDRECEPTOR;
78 PRINIS: PR00498; STEROIDRECEPTOR;
79 PRINIS: PR00498; STEROIDRECEPTOR;
80 PRINIS: PR00498; STEROIDRECEPTOR;
81 PRINIS: PR00498; STEROIDRECEPTOR;
82 PRINIS: PR00498; STEROIDRECEPTOR;
83 PRINIS: PR00498; STEROIDRECEPTOR;
84 PRINIS: PR00498; STEROIDRECEPTOR;
85 PRINIS: PR00498; STEROIDRECEPTOR;
86 PRINIS: PR00498; STEROIDRECEPTOR;
87 PRINIS: PR00498; STEROIDRECEPTOR;
88 PRINIS: PR00498; STEROIDRECEPTOR;
89 PRINIS: PR00498; STEROIDRECEPTOR;
90 PRINIS: PR00498; STEROIDRECEPTOR;
91 PRINIS: PR00498; STEROIDRECEPTOR;
92 PRINIS: PR00498; STEROIDRECEPTOR;
93 PRINIS: PR00498; STEROIDRECEPTOR;
94 PRINIS: PR00498; STEROIDRECEPTOR;
95 PRINIS: PR00498; STEROIDRECEPTOR;
96 PRINIS: PR00498; STEROIDRECEPTOR;
97 PRINIS: PR00498; STEROIDRECEPTOR;
98 PRINIS: PR00498; STEROIDRECEPTOR;
99 PRINIS: PR00498; STEROIDRECEPTOR;
100 PRINIS: PR00498; STEROIDRECEPTOR;

```

```

00 1. FOR THE BINDING ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
01 ER-ALPHA, AND ACTIVATES EXPRESSION OF ESTROGEN-RESPONSE ELEMENTS.
02 ESTROGEN-RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
03 1. SUBUNIT BINDS DNA AS A HOMO-DIMER. CAN FORM A HETERODIMER WITH ER-
04 ALPHA.
05 1. SUBCELLULAR LOCATION: NUCLEAR.
06 1. TISSUE SPECIFICITY: LIVER.
07 1. INDUCTION: BY 17-BETA-ESTRADIOL.
08 1. DOMAIN: COMPOSED OF THREE DOMAINS, A MODULATING N-TERMINAL DOMAIN,
09 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
10 1. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
11 NR3 SUPERFAMILY.
12
13 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
14 between the Swiss Institute of Bioinformatics and the EMBL outstation
15 at the European Bioinformatics Institute. There are no restrictions on its
16 use by non-profit institutions as long as its content is in no way
17 modified and this statement is not removed, usage by and for commercial
18 entities requires a license agreement (see http://www.ebi.ac.uk/Ensembl/
19 or send an email to license@ebi.ac.uk).
20
21 EMBL: AF003456; UAA19851.1;
22 HSSP: P04372; H0CP.
23 InterPro: IPR000824;
24 InterPro: IPR001628;
25 InterPro: IPR001723;
26 InterPro: IPR001729;
27 PRINIS: PR000477; STROIDRECEPTOR;
28 PRINIS: PR00450; VITAMINER;
29 PRINIS: PR00498; STEROIDRECEPTOR;
30 PRINIS: PR00498; STEROIDRECEPTOR;
31 PRINIS: PR00498; STEROIDRECEPTOR;
32 PRINIS: PR00498; STEROIDRECEPTOR;
33 PRINIS: PR00498; STEROIDRECEPTOR;
34 PRINIS: PR00498; STEROIDRECEPTOR;
35 PRINIS: PR00498; STEROIDRECEPTOR;
36 PRINIS: PR00498; STEROIDRECEPTOR;
37 PRINIS: PR00498; STEROIDRECEPTOR;
38 PRINIS: PR00498; STEROIDRECEPTOR;
39 PRINIS: PR00498; STEROIDRECEPTOR;
40 PRINIS: PR00498; STEROIDRECEPTOR;
41 PRINIS: PR00498; STEROIDRECEPTOR;
42 PRINIS: PR00498; STEROIDRECEPTOR;
43 PRINIS: PR00498; STEROIDRECEPTOR;
44 PRINIS: PR00498; STEROIDRECEPTOR;
45 PRINIS: PR00498; STEROIDRECEPTOR;
46 PRINIS: PR00498; STEROIDRECEPTOR;
47 PRINIS: PR00498; STEROIDRECEPTOR;
48 PRINIS: PR00498; STEROIDRECEPTOR;
49 PRINIS: PR00498; STEROIDRECEPTOR;
50 PRINIS: PR00498; STEROIDRECEPTOR;
51 PRINIS: PR00498; STEROIDRECEPTOR;
52 PRINIS: PR00498; STEROIDRECEPTOR;
53 PRINIS: PR00498; STEROIDRECEPTOR;
54 PRINIS: PR00498; STEROIDRECEPTOR;
55 PRINIS: PR00498; STEROIDRECEPTOR;
56 PRINIS: PR00498; STEROIDRECEPTOR;
57 PRINIS: PR00498; STEROIDRECEPTOR;
58 PRINIS: PR00498; STEROIDRECEPTOR;
59 PRINIS: PR00498; STEROIDRECEPTOR;
60 PRINIS: PR00498; STEROIDRECEPTOR;
61 PRINIS: PR00498; STEROIDRECEPTOR;
62 PRINIS: PR00498; STEROIDRECEPTOR;
63 PRINIS: PR00498; STEROIDRECEPTOR;
64 PRINIS: PR00498; STEROIDRECEPTOR;
65 PRINIS: PR00498; STEROIDRECEPTOR;
66 PRINIS: PR00498; STEROIDRECEPTOR;
67 PRINIS: PR00498; STEROIDRECEPTOR;
68 PRINIS: PR00498; STEROIDRECEPTOR;
69 PRINIS: PR00498; STEROIDRECEPTOR;
70 PRINIS: PR00498; STEROIDRECEPTOR;
71 PRINIS: PR00498; STEROIDRECEPTOR;
72 PRINIS: PR00498; STEROIDRECEPTOR;
73 PRINIS: PR00498; STEROIDRECEPTOR;
74 PRINIS: PR00498; STEROIDRECEPTOR;
75 PRINIS: PR00498; STEROIDRECEPTOR;
76 PRINIS: PR00498; STEROIDRECEPTOR;
77 PRINIS: PR00498; STEROIDRECEPTOR;
78 PRINIS: PR00498; STEROIDRECEPTOR;
79 PRINIS: PR00498; STEROIDRECEPTOR;
80 PRINIS: PR00498; STEROIDRECEPTOR;
81 PRINIS: PR00498; STEROIDRECEPTOR;
82 PRINIS: PR00498; STEROIDRECEPTOR;
83 PRINIS: PR00498; STEROIDRECEPTOR;
84 PRINIS: PR00498; STEROIDRECEPTOR;
85 PRINIS: PR00498; STEROIDRECEPTOR;
86 PRINIS: PR00498; STEROIDRECEPTOR;
87 PRINIS: PR00498; STEROIDRECEPTOR;
88 PRINIS: PR00498; STEROIDRECEPTOR;
89 PRINIS: PR00498; STEROIDRECEPTOR;
90 PRINIS: PR00498; STEROIDRECEPTOR;
91 PRINIS: PR00498; STEROIDRECEPTOR;
92 PRINIS: PR00498; STEROIDRECEPTOR;
93 PRINIS: PR00498; STEROIDRECEPTOR;
94 PRINIS: PR00498; STEROIDRECEPTOR;
95 PRINIS: PR00498; STEROIDRECEPTOR;
96 PRINIS: PR00498; STEROIDRECEPTOR;
97 PRINIS: PR00498; STEROIDRECEPTOR;
98 PRINIS: PR00498; STEROIDRECEPTOR;
99 PRINIS: PR00498; STEROIDRECEPTOR;
100 PRINIS: PR00498; STEROIDRECEPTOR;

```

```

CC -1- SUBUNIT BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
CC ALPHA (BY SIMILARITY) NUCLEAR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF061269; AACD26921.1; -.
CC HSSP: P03372; HCP.
CC InterPro: IPR000924; -.
CC DR InterPro: IPR001628; STROIDFINGER.
CC DR PRINTS: PRO00047; STROIDFINGER.
CC DR PRINTS: PRO0350; VITAMINE.
CC DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger; Steroid-binding.
CC KW Domain; Steroid-binding.
CC KW Domain; 12 169 MODULATING.
CC FT DNA_BIND 170 232 NUCLEAR PROTEIN-TYPE
CC FT ZN_FING 170 190 CA-TYPE
CC FT ZN_FING 206 240 CA-TYPE.
CC FT DOMAIN 216 568 STEROID-BINDING.
CC FT SEQUENCE 568 AA; 63539 MW; 2D678D775D876F9 CRC64;
CC -----
CC Query Match: 25.5%; Score 61.5; DB 1; Length 568;
CC Best Local Similarity: 43.4%; Prod No. 2.9;
CC Matches: 14; Conservative: 5; Mismatches: 12; Indels 1; Gaps 1;
CC -----
CC 12 STSSVNCSSSTLPLEMGSVYIFSSVVDSHHY 43
CC 1) : : : : 11 : : 111111 : : : :
CC DB 38 SP-TFSSSDPLVESHPICLIPSPYDIDCHDF 68
CC -----
CC RESULT 11
CC YHFL_YEAST
CC ID YHFL_YEAST STANDARD: PRT: 668 AA.
CC EC P4217;
CC DT 01-FEB-1994 (Rel. 28, (created)
CC DT 01-FEB-1994 (Rel. 28, last sequence update)
CC DT 01-OCT-2000 (Rel. 40, last annotation update)
CC DE HYPOHETICAL 73.8 KDA PROTEIN IN SAS3-SPEC17 INTERFERING PRION
CC GN YH051C OR YH0516 OR YH0506.
CC OS Saccharomyces cerevisiae (Baker's yeast)
CC EC Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes;
CC Saccharomycetales, Saccharomycetaceae, Saccharomyces.
CC NCBI_TaxID:4932;
CC OR 111
CC RA SEQUENCE FROM N.A.
CC RC STRAIN-528BC;
CC RA Melnik-420C-Ly, Falmud R14187;
CC RA Scheraga B., el Bakoury M., Vetterli's F., Dubois E., Messerogy F.;
CC RT Sequencing and functional analysis of a 32,560 bp segment of the
CC left arm of yeast chromosome II. Identification of 26 open reading
CC frames, including the KIF1 and SEC17 genes.
CC RT Yeast 9:155-1371(1993).
CC KL -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RPM).
CC -1- SIMILARITY: SOME, TO S. POMBE SPAC12C12.03.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@sb.stj.ch).
CC
DR      EMBL_223261; CAA80795.1; -
DR      EMBL_225812; CAA84871.1; -
DR      PIR_539636; S39836.
DR      PIR_537337; S37337.
DR      SDB_Sequid47; YH1051C.
DR      InterPro: IPR00504; -.
DR      Pfam: PF00076; Tm; 1.
DR      ProSITE: PS0102; RRM_RNP_1; 1.
DR      PROSITE: PS00303; RRM_RNP_1; 1.
DR      Hypothetical protein; RNA-binding.
KW      DOMAIN 85 163 163 MM. LAMFSL2664 FCIB CRG64.
SQ      SEQUENCE 668 AA; 7377 MM. LAMFSL2664 FCIB CRG64.

Query Match          24.3%; Score 58.5; DB 1; Length 668;
Host Local Similarity 40.0%; Prod. No. 8.1;
Matches 14; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

OY      9 SUNSPSYNCSOSTIPLE-HGSTYTPTSSYDISHHE 42
           :||| | | | | | | | | | | | | | | | |
Db       585 NINQPSANQSOLLPOHTNGSHNSNYGSYNDE 619

RESULT 12
AMAL_PLACH ID AMAL_PLACH STANDARD: PRG: 558 AA.
AC P16445;
DT 01-APR-1990 (rel. 15, Created)
BT 01-APR-1990 (rel. 15, Last sequence update)
DT 01-NOV-1997 (rel. 35, Last annotation update)
DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MERZOITITE SURFACE ANTIGEN).
GN AMA-1.
OS Plasmodium chabaudi.
OU Fukuiyoda, Alzeolactz, Apicomplexa; Haemosporida; Plasmodium.
OA NCBI_TaxID:5825;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DK;
RX MEDLINE=90114335; PubMed=2608101;
RA Marshall V.M., Peterson M.G., Lew A.M., Kemp D.J.;
RT "Structure of the apical membrane antigen 1 (AMA-1) of Plasmodium
   chabaudi."
RL Mol. Biochem. Parasitol. 37:281-284(1989).
[2]
RP SEQUENCE OF 38-377 FROM N.A.
RC STRAIN-DK;
RX MEDLINE=96333375; PubMed=8757869;
RA Crewther P.E., Matthew M.L., Flegel R.H., Anders K.P.;
RT "Protective immune responses to apical membrane antigen 1 of
   Plasmodium chabaudi involve recognition of strain-specific
   epitopes."
RI Infect. Immun. 64 3310 3317(1996);
CC -1- INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC ! SIMILARITY: STRONG TO AMA-1/PfPR3 FROM P.FALCIPARUM, P.FRAGILE AND
   TO PK66 FROM P.KNOWLESII.
-----
CC This Swiss Prot entry is copyright. It is prepared through a collaboration
CC between The Swiss Institute of Bioinformatics and the EMBL contribution.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see http://www.ispi.ch/homeowner/
CC or send an email to license@sb.stj.ch).
CC
DR      EMBL_M25248; AAA90929.1; -
DR      EMBL_A08270; CAA00765.1; -
DR      EMBL_049745; AAB36511.1; -
DR      PIR_AA4964; A44964.
KW      Majorita. Signal; Transmembrane; Antigen; Glycoprotein.
FT      SIGNAL 1 2 POTENTIAL.
```

```

CC      or send an email to license@sb.stb.ch
CC
DR      EMBL: 223261; CAA8795.1; -
DR      EMBL: 225812; CAA8487.1; -
DR      PIR: S39636; S39836.
DR      PIR: S37337; S3737.
DR      SDB: Ssqm0147; YH1051C.
DR      InterPro: IPR00504; -.
DR      Pfam: PF0076; Trm; 1.
DR      ProSITE: PS0102; RRM_RNP_1; 1.
DR      ProSITE: PS0030; RRM_RNP_1; 1.
DR      Hypothetical protein; RNA-binding.
KW      DOMAIN 85 163 163 MM. EMBL:2664563 CIG64.
SQ      SEQUENCE 668 AA; 7377 MM. EMBL:2664563 CIG64.
CC
Query Match 24.3%; Score 58.5; DB 1; Length 668;
Query Local Similarity 40.0%; Pred. No. 8.1;
Matches 14; Conservative 5; Mismatches 15; Indels 1; Gaps 1;
CC
DB 585 NINSPSSANSGQLLPHTNGSHNSNGSYSDH 619
CC
CC      RESULT 12
CC      AMAL_PLACH STANDARD: PRT; 558 AA.
CC      ID AMAL_PLACH
CC      AC P16445;
CC      DT 01-APR-1990 (rel. 15, Created)
CC      DT 01-APR-1990 (rel. 15, Last sequence update)
CC      DT 01-NOV-1997 (rel. 35, Last annotation update)
CC      DE APICAL MEMBRANE ANTIGEN 1 PRECURSOR (MEROZOITE SURFACE ANTIGEN).
CC      GN AMA-1.
CC      OS Plasmodium chabaudi.
CC      ON Fkayiyota, Alveolates, Apicomplexa; Haemosporida; Plasmodium.
CC      OA NCBI_TaxID:5825;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN-DS;
CC      RX MEDLINE:90114335; PubMed:2608101;
CC      RA Marshall V.M., Peterson M.G., Lew A.M., Kemp D.J.;
CC      RT "Structure of the apical membrane antigen 1 (AMA-1) of Plasmodium
CC      chabaudi."
CC      RL Mol. Biochem. Parasitol. 37:281-284(1989).
CC      [2]
CC      RP SEQUENCE OF 38-377 FROM N.A.
CC      RC STRAIN-DK;
CC      RX MEDLINE:96333375; PubMed:8757869;
CC      RA Crewther P.E., Matthew M.L., Flegel R.H., Anders R.P.;
CC      RT "Protective immune responses to apical membrane antigen 1 of
CC      Plasmodium chabaudi involve recognition of strain-specific
CC      epitopes."
CC      RL Infect. Immun. 64 3310-3317(1996).
CC      CC -1- INFECTION INVOLVED IN PARASITE INVASION OF ERYTHROCYTES.
CC      CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC      CC -1- SIMILARITY: STRONG TO AMA-1/PEP3 FROM P.FALCIPARUM, P.FRAGILE AND
CC      TO P666 FROM P.KNOWLESII.
CC
CC      -----
CC      This Swiss Prot entry is copyright. It is prepared through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL collaboration.
CC      The European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed, usage by and for commercial
CC      entities requires a license agreement (info@ebi.scrib.ch/
CC      license@sb.stb.ch).
CC      or send an email to license@sb.stb.ch.
CC
CC      EMBL: M25248; AAA9029.1; -
CC      EMBL: A08270; CAA00765.1; -
CC      EMBL: 049745; AAB36511.1; -
CC      PIR: A44964; A44964.
CC      Material: Signal; Transmembrane; Antigen; Glycoprotein.
CC      SIGNAL 1 2 POTENTIAL.
CC

```



```

RL Curr. Microbiol. 37:356-358(1998).
CC -1- FUNCTION. THIS PROTEIN IS CONSIDERED AS A BIFUNCTIONAL ENZYME,
CC POSSESSING TWO ACTIVE SITES. ONE AN ALCOHOL DEHYDROGENASE AND
CC THE OTHER AN ALDEHYDE DEHYDROGENASE (BY SIMILARITY)
CC -1- CATALYTIC ACTIVITY: L-HISTIDINOL + 2 NAD(+) -> L-HISTIDINE +
CC 2 NADH.
CC -1- PATHWAY: TENTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: TO OTHER PROKARYOTIC, FUNGAL AND PLANTS HDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF067228; AAC97355.1; -.
DR InterPro: IPR001692; -.
DR InterPro: IPR001917; -.
DR Pfam: PF00815; Histidinol_dh; 1.
DR PRINTS: PR00083; HDHDEHGNASE.
DR PROSITE: PS00611; HISOL_DEHYDROGNASE; FALSE_NEG.
DR Histidine biosynthesis: Multistep histidinol conversion: Cofactor: NAD.
DR SPOUENGL 435 AA; 4737 SW; CD233872.DEDDED CKC64;
SQ

```

```

Query Match          23.2%; Score 56; DB 1; Length 435;
Post Local Similarity 41.4%; Pred. No. 10;
Matches 12; Conservative 2; Mismatches 15; Indels 0; Gaps 0;
UY 6 SPSSLSPPSYNCOSIPLHGSIVIPS 34
  | : : | | | | | | | | | |
Db 106 SPIDLETGVGVCQIVLEPNSVGIYIPS 134

```

Search completed: April 19, 2001, 10:35:35
Job time: 14774 sec







405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Query Match
Best Local Similarity
Matches 431
Conservative
1 MINKSPSSLSINSPSSVNFQSSILPHFQSTVTPSSVYDSHBYHPAMTFYSPAVMNSYTPS 60
1 MINKSPSSLSINSPSSVNFQSSILPHFQSTVTPSSVYDSHBYHPAMTFYSPAVMNSYTPS 79

405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Query Match
Best Local Similarity
Matches 431
Conservative
1 MINKSPSSLSINSPSSVNFQSSILPHFQSTVTPSSVYDSHBYHPAMTFYSPAVMNSYTPS 60
1 MINKSPSSLSINSPSSVNFQSSILPHFQSTVTPSSVYDSHBYHPAMTFYSPAVMNSYTPS 79

